

PATENT COOPERATION TREATY

PCT

NOTIFICATION OF ELECTION

(PCT Rule 61.2)

From the INTERNATIONAL BUREAU

To:

Assistant Commissioner for Patents
United States Patent and Trademark
Office
Box PCT
Washington, D.C.20231
ETATS-UNIS D'AMERIQUE

in its capacity as elected Office

Date of mailing (day/month/year) 08 June 2000 (08.06.00)	Applicant's or agent's file reference 53371-56822
International application No. PCT/SE99/01784	Priority date (day/month/year) 06 October 1998 (06.10.98)
International filing date (day/month/year) 06 October 1999 (06.10.99)	
Applicant ZAPHIROPOULOS, Peter, G. et al	

1. The designated Office is hereby notified of its election made:

☒ in the demand filed with the International Preliminary Examining Authority on:
05 May 2000 (05.05.00)

☐ in a notice effecting later election filed with the International Bureau on:

2. The election ☒ was

☐ was not

made before the expiration of 19 months from the priority date or, where Rule 32 applies, within the time limit under Rule 32.2(b).

The International Bureau of WIPO 34, chemin des Colombettes 1211 Geneva 20, Switzerland Facsimile No.: (41-22) 740.14.35	Authorized officer F. Baechler Telephone No.: (41-22) 338.83.38
---	---

REC'D 24 JAN 2001

INTERNATIONAL PRELIMINARY EXAMINATION REPORT

(PCT Article 36 and Rule 70)

15

Applicant's or agent's file reference 53371-56822		FOR FURTHER ACTION See Notification of Transmittal of International Preliminary Examination Report (Form PCT/IPEA/416)	
International application No. PCT/SE99/01784	International filing date (day/month/year) 06.10.1999	Priority date (day/month/year) 06.10.1998	
International Patent Classification (IPC) or national classification and IPC ₇ A61K 48/00, C07K 14/705, A61K 38/17, G01N 33/68			
Applicant Karolinska Innovations AB et al			

1. This international preliminary examination report has been prepared by this International Preliminary Examining Authority and is transmitted to the applicant according to Article 36.
2. This REPORT consists of a total of 6 sheets, including this cover sheet.

☒ This report is also accompanied by ANNEXES, i.e., sheets of the description, claims and/or drawings which have been amended and are the basis for this report and/or sheets containing rectifications made before this Authority (see Rule 70.16 and Section 607 of the Administrative Instructions under the PCT).

These annexes consist of a total of 2 sheets.

3. This report contains indications relating to the following items:

- I ☒ Basis of the report
- II ☐ Priority
- III ☒ Non-establishment of opinion with regard to novelty, inventive step and industrial applicability
- IV ☐ Lack of unity of invention
- V ☒ Reasoned statement under Article 35(2) with regard to novelty, inventive step or industrial applicability; citations and explanations supporting such statement
- VI ☐ Certain documents cited
- VII ☐ Certain defects in the international application
- VIII ☐ Certain observations on the international application

Date of submission of the demand 05.05.2000	Date of completion of this report 12.01.2001
Name and mailing address of the IPEA/SE Patent- och registreringsverket Box 5055 S-102 42 STOCKHOLM Facsimile No. 08-667 72 88	Authorized officer Patrick Andersson/EÖ Telephone No. 08-782 25 00

INTERNATIONAL PRELIMINARY EXAMINATION REPORT

International application No.

PCT/SE99/01784

I. Basis of the report

1. With regard to the **elements** of the international application:*

- ☐ the international application as originally filed
- ☒ the description:
pages 1-23 , as originally filed
pages _____ , filed with the demand
pages _____ , filed with the letter of _____
- ☒ the claims:
pages _____ , as originally filed
pages _____ , as amended (together with any statement) under article 19
pages _____ , filed with the demand
pages 1, 2 , filed with the letter of 15.12.2000
- ☐ the drawings:
pages _____ , as originally filed
pages _____ , filed with the demand
pages _____ , filed with the letter of _____
- ☐ the sequence listing part of the description:
pages _____ , as originally filed
pages _____ , filed with the demand
pages _____ , filed with the letter of _____

2. With regard to the **language**, all the elements marked above were available or furnished to this Authority in the language in which the international application was filed, unless otherwise indicated under this item.

These elements were available or furnished to this Authority in the following language _____ which is:

- ☐ the language of a translation furnished for the purposes of international search (under Rule 23.1(b)).
- ☐ the language of publication of the international application (under Rule 48.3(b)).
- ☐ the language of the translation furnished for the purposes of international preliminary examination (under Rules 55.2 and/or 55.3).

3. With regard to any **nucleotide and/or amino acid sequence** disclosed in the international application, the international preliminary examination was carried out on the basis of the sequence listing:

- ☐ contained in the international application in written form.
- ☐ filed together with the international application in computer readable form.
- ☐ furnished subsequently to this Authority in written form.
- ☐ furnished subsequently to this Authority in computer readable form.
- ☐ The statement that the subsequently furnished written sequence listing does not go beyond the disclosure in the international application as filed has been furnished.
- ☐ The statement that the information recorded in computer readable form is identical to the written sequence listing has been furnished.

4. ☐ The amendments have resulted in the cancellation of:

- ☐ the description, pages _____
- ☐ the claims, Nos. _____
- ☐ the drawings, sheet/fig _____

5. ☐ This report has been established as if (some of) the amendments had not been made, since they have been considered to go beyond the disclosure as filed, as indicated in the Supplemental Box (Rule 70.2 (c)).**

* Replacement sheets which have been furnished to the receiving Office in response to an invitation under Article 14 are referred to in this report as "originally filed" and are annexed to this report since they do not contain amendments (Rules 70.16 and 70.17).

** Any replacement sheet containing such amendments must be referred to under item I and annexed to this report.

INTERNATIONAL PRELIMINARY EXAMINATION REPORT

International application No.

PCT/SE99/01784

III. Non-establishment of opinion with regard to novelty, inventive step and industrial applicability

1. The questions whether the claimed invention appears to be novel, to involve an inventive step (to be non obvious), or to be industrially applicable have not been examined in respect of:

☐ the entire international application,

☒ claims Nos. 7 (partially), 8-10 and 16 completely

because:

☒ the said international application, or the said claims Nos. 7 (partially) and 16 completely relate to the following subject matter which does not require an international preliminary examination (*specify*):

Claims 7 partially 16 completely relate to methods of treatment of the human or animal body by therapy/ diagnostic methods practised on the human or animal body See PCT Rule. 67.1. (iv).

☒ the description, claims or drawings (*indicate particular elements below*) or said claims Nos. 8-10 are so unclear that no meaningful opinion could be formed (*specify*):

Claim 8 relates to a screening method, it is not clear from this claim or the description how the polypeptide is used to achieve the screening. Claim 9 relates to a method of synthesis a modified drug, it is not clear from the description or the claim how the protein is used. Claim 10 relates to a drug identified by any of the methods in 8 or 9. Thus, claims 8-10 is considered to fail to comply with PCT-article 6 and have not been searched. Moreover, claim 15 concerns a drug per se, the drug may be a known substance.

.../...

☐ the claims, or said claims Nos. _____ are so inadequately supported by the description that no meaningful opinion could be formed.

☐ no international search report has been established for said claims Nos. _____

2. A meaningful international preliminary examination cannot be carried out due to the failure of the nucleotide and/or amino acid sequence listing to comply with the standard provided for in Annex C of the Administrative Instructions:

☐ the written form has not been furnished or does not comply with the standard.

☐ the computer readable form has not been furnished or does not comply with the standard.

INTERNATIONAL PRELIMINARY EXAMINATION REPORT

International application No.

PCT/SE99/01784

Supplemental Box

(To be used when the space in any of the preceding boxes is not sufficient)

Continuation of: III

The description or the claims does not give any indication on that any unknown substance(s) that are preferred. For this reason, a search or a statement concerning novelty and an inventive step can not be made for claim 10.

INTERNATIONAL PRELIMINARY EXAMINATION REPORT

International application No.

PCT/SE99/01784

V. Reasoned statement under Article 35(2) with regard to novelty, inventive step or industrial applicability; citations and explanations supporting such statement**1. Statement**

Novelty (N)	Claims	<u>5-7, 11-15</u>	YES
	Claims	<u>1-4, 17</u>	NO
Inventive step (IS)	Claims		YES
	Claims	<u>1-7, 11-15, 17</u>	NO
Industrial applicability (IA)	Claims	<u>1-7, 11-15, 17</u>	YES
	Claims		NO

2. Citations and explanations (Rule 70.7)

The claimed invention relates to nucleic acids, proteins participating in the human PTCH/SHH pathway as well as related vectors, cells, antibodies, methods and uses.

The following document is considered relevant:

D1) Motoyama J et al., "'Ptch2, a second mouse Patched gene is co-expressed with sonic hedgehog"

D1 discloses a mouse patched gene and its amino acid sequence having 91% homology in an 1140 amino acid overlap. The expression is analysed with RNA probes. In the present description a comparison between the sequence of D1 and the sequence claimed is made; the differences between the sequences are pointed out and it is speculated what effect these differences may have. However, the difference in function between the two has not been shown. Therefore, the expression "a variant thereof capable of participating in the human PTCH/SHH pathway" is unclear since a person skilled in the art get no information on what is covered e.g. by experiments showing unique functions of the claimed protein. The protein/nucleic acid of D1 would interact(i.e. be capable of participating) in some way also in the human pathway and can therefore be regarded as a variant of SEQ ID NO: 1. Thus, the invention according claims 1-4 and 17 lacks novelty.

.../...

INTERNATIONAL PRELIMINARY EXAMINATION REPORT

International application No.

PCT/SE99/01784

Supplemental Box

(To be used when the space in any of the preceding boxes is not sufficient)

Continuation of: V

Nucleic acids hybridising to a nucleic acid, vectors, recombinant cells, antibodies, kits are all features obvious to a person skilled in the art; whether claims 11-15 involves an inventive step depends on if the claims the relate to are inventive.

Thus the invention according to claims 11-15 is novel and industrially applicable, but is not considered to involve an inventive step.

D1 discloses that the patched gene is involved in cancer, that the gene or its product can be used in some way in a medicament or diagnoses is considered obvious to a person skilled in the art. The present application seems not disclose any specific uses in medicament or diagnoses (e.g. a link between a mutation in the patched and a medical state) that goes beyond what a person skilled in the art would expect in view of D1. Thus the invention according to claims 5-7 is novel but not considered to involve an inventive step.

CLAIMS

1. An isolated human protein or an analogue or variant thereof capable of participating in the human PTCH/SHH pathway during embryonic development and/or carcinogenesis comprising at least about 1040 amino acids of SEQ ID NO: 1.
- 5 2. A protein according to claim 1, which is essentially comprised of SEQ ID NO: 1.
3. A nucleic acid encoding a protein according to any one of claims 1 and 2.
4. A nucleic acid encoding a protein according to claim 3 comprising at least about 3094 bases of SEQ ID NO: 2.
- 10 5. An isolated genomic nucleic acid comprising parts or all of SEQ ID NO: 5.
6. An isolated variant of the nucleic acid according to claim 5.
7. A nucleic acid having the sequence of a splicing variant selected from the group of variants defined in Figure 2B and SEQ ID NO:3 and SEQ ID NO:4.
8. An isolated nucleic acid capable of specifically hybridising to a nucleic acid according to any one of claims 3-7.
- 15 9. A protein or polypeptide encoded by a nucleic acid according to claim 7.
10. A protein according to claim 1 or 2, a nucleic acid according to any one of claims 2-8 or or a polypeptide according to claim 9, for use as a medicament.
11. Use of a protein according to claim 1 or 2, a nucleic acid according to any one of claims 2-8 or a polypeptide according to claim 9 in the manufacture of a medicament for the treatment of a condition involving tumors, such as BSS.
- 20 12. A method of in vitro or in vivo diagnosis, wherein a protein according to claim 1 or 2, a nucleic acid according to any one of claims 2-8 or a polypeptide according to claim 9 is used.
- 25 13. A method of screening wherein a library of suitable candidate compounds is screened for modified drugs using a protein according to claim 1 or 2 or a polypeptide according to claim 9 as a lead compound.
14. A method of synthesis of a modified drug, wherein a protein according to claim 1 or 2, or a protein or polypeptide according to claim 9, is used.
- 30 15. A modified drug identified by the method according to claim 13 or synthesized according to claim 14.

16. A vector comprising a nucleic acid according to any one of claim 3-8. —

17. A recombinant cell comprising a vector according to claim 16.

18. An antibody which specifically binds to a protein or polypeptide according to claim 1, 2 or 9.

5 19. A recombinant cell expressing an antibody according to claim 18.

20. A kit for the detection of a human PTCH2 gene or polypeptide comprising in a container a molecule selected from the group consisting of a nucleic acid according to any one of claims 3-8, a polypeptide or protein according to claim 1, 2 or 9 or an antibody according to claim 18.

10 21. Use of a nucleic acid according to any one of claims 3-7 in gene therapy.

22. Use of nucleic acid according to any one of claims 3-7 as a probe, a primer or a diagnostic reagent.

PCTWORLD INTELLECTUAL PROPERTY ORGANIZATION
International Bureau

INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(51) International Patent Classification 7 : A61K 48/00, C07K 14/705, A61K 38/17, G01N 33/68		A1	(11) International Publication Number: WO 00/20037
			(43) International Publication Date: 13 April 2000 (13.04.00)
(21) International Application Number: PCT/SE99/01784		(81) Designated States: AE, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, CA, CH, CN, CR, CU, CZ, DE, DK, DM, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MD, MG, MK, MN, MW, MX, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, TZ, UA, UG, US, UZ, VN, YU, ZA, ZW, ARIPO patent (GH, GM, KE, LS, MW, SD, SL, SZ, TZ, UG, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GW, ML, MR, NE, SN, TD, TG).	
(22) International Filing Date: 6 October 1999 (06.10.99)			
(30) Priority Data: 9803393-9A ✓ 6 October 1998 (06.10.98) SE			
(71) Applicants (for all designated States except US): KAROLIN-SKA INNOVATIONS AB [SE/SE]; S-171 77 Stockholm (SE). PHARMACIA & UPJOHN COMPANY [US/US]; 301 Henrietta Street, Kalamazoo, MI 49001 (US).			
(72) Inventors; and (75) Inventors/Applicants (for US only): ZAPHIROPOULOS, Peter, G. [GR/SE]; Tullinge Strand 96, S-146 54 Tullinge (SE). UNDÉN, Anne, Birgitte [NO/SE]; Gandviksvägen 3, S-181 62 Djursholm (SE). TOFTGÅRD, Rune [SE/SE]; Sättagårdsvägen 209, S-127 36 Skärholmen (SE). RAHNAMA, Fahimeh [SE/SE]; Stadshagsvägen 18 3 tr, S-112 50 Stockholm (SE). HOLLINGSWORTH, Robert, E. [US/US]; Glaxo Inc., 5 Moore Drive, Research Triangle Park, NC 27709 (US).		Published <i>With international search report. Before the expiration of the time limit for amending the claims and to be republished in the event of the receipt of amendments.</i>	
(74) Agents: BERG, S., A. et al.; Albihns Patentbyrå Stockholm AB, P.O. Box 5581, S-114 85 Stockholm (SE).			
(54) Title: A NOVEL COMPONENT IN THE HEDGEHOG SIGNALLING PATHWAY			
(57) Abstract <p>The present invention relates to a novel human patched-like gene (PTCH2), which for the first time has been cloned and sequenced. Several alternatively spliced mRNA forms of PTCH2 have been identified, including transcripts lacking segments thought to be involved in sonic hedgehog (SHH) binding and mRNAs with differentially defined 3' terminal exons. Further, the invention also relates to the protein encoded by the present PTCH2 as well as to functional analogues and variants thereof.</p>			

FOR THE PURPOSES OF INFORMATION ONLY

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

AL	Albania	ES	Spain	LS	Lesotho	SI	Slovenia
AM	Armenia	FI	Finland	LT	Lithuania	SK	Slovakia
AT	Austria	FR	France	LU	Luxembourg	SN	Senegal
AU	Australia	GA	Gabon	LV	Latvia	SZ	Swaziland
AZ	Azerbaijan	GB	United Kingdom	MC	Monaco	TD	Chad
BA	Bosnia and Herzegovina	GE	Georgia	MD	Republic of Moldova	TG	Togo
BB	Barbados	GH	Ghana	MG	Madagascar	TJ	Tajikistan
BE	Belgium	GN	Guinea	MK	The former Yugoslav Republic of Macedonia	TM	Turkmenistan
BF	Burkina Faso	GR	Greece			TR	Turkey
BG	Bulgaria	HU	Hungary	ML	Mali	TT	Trinidad and Tobago
BJ	Benin	IE	Ireland	MN	Mongolia	UA	Ukraine
BR	Brazil	IL	Israel	MR	Mauritania	UG	Uganda
BY	Belarus	IS	Iceland	MW	Malawi	US	United States of America
CA	Canada	IT	Italy	MX	Mexico	UZ	Uzbekistan
CF	Central African Republic	JP	Japan	NE	Niger	VN	Viet Nam
CG	Congo	KE	Kenya	NL	Netherlands	YU	Yugoslavia
CH	Switzerland	KG	Kyrgyzstan	NO	Norway	ZW	Zimbabwe
CI	Côte d'Ivoire	KP	Democratic People's Republic of Korea	NZ	New Zealand		
CM	Cameroon		Republic of Korea	PL	Poland		
CN	China	KR	Republic of Korea	PT	Portugal		
CU	Cuba	KZ	Kazakstan	RO	Romania		
CZ	Czech Republic	LC	Saint Lucia	RU	Russian Federation		
DE	Germany	LI	Liechtenstein	SD	Sudan		
DK	Denmark	LK	Sri Lanka	SE	Sweden		
EE	Estonia	LR	Liberia	SG	Singapore		

A NOVEL COMPONENT IN THE HEDGEHOG SIGNALLING PATHWAY

Technical field

The present invention relates to novel molecules, such as proteins, polypeptides and nucleotides, involved in the hedgehog signalling pathway with putative involvement in embryonic development and carcinogenesis. The invention also relates to various novel advantageous uses of the molecules according to the invention, e.g. in diagnosis and therapy.

Background

In the study of the development of cells, fruit flies have extensively been used as a model, as they are less complex than mammalian cells.

Pattern formation takes place through a series of logical steps, reiterated many times during the development of an organism. Viewed from a broader evolutionary perspective, across species, the same sort of reiterative pattern formations are seen. The central dogma of pattern formation has been described (Lawrence and Struhl, 1996). Three interlocking and overlapping steps are defined. Firstly, positional information in the form of morphogen gradients allocate cells into non-overlapping sets, each set founding a compartment. Secondly, each of these compartments acquire a genetic address, as a result of the function of active "selector" genes, that specify cell fate within a compartment and also instruct cells and their descendents how to communicate with cells in neighboring compartments. The third step involves interactions between cells in adjacent compartments, initiating new morphogen gradients, which directly organize the pattern.

Taking these steps in greater detail, one finds the first step in patterning to be the definition of sets of cells in each primordium. Cells are allocated according to their positions with respect to both dorsoventral and anterior/posterior axes by morphogen gradients. Allocation of cells in the dorsoventral axis constitutes the germ layers, such as mesoderm or neurectoderm.

In segmentation, the second step (the specification of cell fate in each compartment) is carried out by the gene *engrailed* and elements of the bithorax complex. *Engrailed* defines anterior and posterior compartments both in segmentation and in limb specification.

5

10

15

20

25

30

The third step in pattern formation, secretion of morphogens, functions to differentiate patterns within compartments (and thereby establish segment polarity). Initially, all cells within a compartment are equipotent, but they become diversified to form pattern. Pattern formation depends on gradients of morphogens, gradients initiated along compartment boundaries. Such gradients are established by a short-range signal induced in all the cells of the compartment in which the above mentioned selector gene *engrailed* is active. For segment polarity, this signal is Hedgehog. In the adjacent compartment the selector gene is inactive, ensuring that the cells are sensitive to the signal. The Hedgehog signal range is probably only a few rows of cells wide; responding cells become a linear source of a long-range morphogen, that diffuses outward in all directions. There are three known Hedgehogs, Sonic (SHH), Indian (IHH) and Desert (DHH). The proteins they encode can substitute each for each other, but in wildtype animals, their distinct distributions result in unique activities. SHH controls the polarity of limb growth, directs the development of neurons in the ventral neural tube and patterns somities. IHH controls endochondral bone development and DHH is necessary for spermiogenesis. Vertebrate hedgehog genes are expressed in many other tissues, including the peripheral nervous system, brain, lung, liver, kidney, tooth primordia, genitalia and hindgut and foregut endoderm.

Thus, segment polarity genes have been identified in flies as mutations, which change the pattern of structures of the body segments. Mutations in these genes cause animals to develop the changed patterns on the surfaces of body segments, the changes affecting the pattern along the head to tail axis. For example, mutations in the gene *patched* cause each body segment to develop without the normal structures in the center of each segment. Instead there is a mirror image of the pattern normally found in the anterior segment. Thus, cells in the center of the segment make the

wrong structures, and point them in the wrong direction with reference to the over all head-to-tail polarity of the animal.

About sixteen genes in the class are known. The encoded proteins include kinases, transcription factors, a cell junction protein, two secreted proteins called wingless (WG) and the above mentioned Hedgehog (HH), a single transmembrane protein called patched (PTC) and some novel proteins not related to any known protein. All of these proteins are believed to work together in signaling pathways that inform cells about their neighbors in order to set cell fates and polarities.

PTC has been proposed as a receptor for HH protein based on genetic experiments in flies. A model for the relationship is that PTC acts through a largely unknown pathway to inactivate both its own transcription and the transcription of the *wingless* segment polarity gene. This model proposes that HH protein, secreted from adjacent cells, binds to the PTC receptor, inactivates it and thereby prevents PTC from turning off its own transcription or that of *wingless*. A number of experiments have shown coordinate events between PTC and HH.

Human *patched* gene (PTCH) was recently identified as the gene responsible for the nevoid basal cell carcinoma syndrome (NBCCS), also known as the Gorlin Syndrome, which is an autosomal dominant disorder that predisposes to both cancer and developmental defects (Gorlin (1995) *Dermatologic Clinics* 13:113-125) characterized by multiple basal cell carcinomas (BCCs), medulloblastomas and ovarian fibromas as well as numerous developmental anomalies (Hahn, H., Wicking, C., Zaphiropoulos, P.G., Gailani, M.R., Shanley, S., Chidambaram, A., Vorechovsky, I., Holmberg, E., Undén, A.B., Gillies, S., Negus, K., Smyth, I., Pressman, C., Lefell, D.J., Gerrard, B., Goldstein, A.M., Dean, M., Toftgård, R., Chenevix-Trench, G., Wainright, B. and Bale, A.E. (1996): "Mutations of the human homolog of *Drosophila* patched in the nevoid basal cell carcinoma syndrome", *Cell* 85, 841-851; and Johnson, R.L., Rothman, A.L., Xie, J., Goodrich, L.V., Bare, J.W., Bonifas, J.M., Quinn, A.G., Myers, R.M., Cox, D.R., Epstein, E.H. Jr and Scott, M.P.

(1996): "Human homolog of patched, a candidate gene for the basal cell nevus syndrome", *Science* **272**, 1668-1671). PTCH codes for a membrane receptor of the autolytically cleaved (protein spliced), amino terminal domain of *sonic hedgehog* (SHH) (Mariago, V., Davey, R.A., Zuo, Y., Cunningham, J.M. and Tabin, C.J. (1996): "Biochemical evidence that patched is the Hedgehog receptor", *Nature* **384**, 176-179; and Stone, D.M., Hynes, M., Armanini, M., Swanson, T.A., Gu, Q., Johnson, R.L., Scott, M.P., Pennica, D., Goddard, A., Phillips, H., Noll, M., Hooper, J.E., de Sauvage, F. and Rosenthal, A. (1996): "The tumor-suppressor gene patched encodes a candidate receptor for Sonic hedgehog", *Nature* **384**, 129-134). In the non-signalling state, PTCH is thought to inhibit the consecutive signalling of another membrane protein, *smoothened* (SMO), however binding of SHH to PTCH relieves this inhibition (Goodrich, L.V., Milenkovic, L., Higgins, K.M. and Scott, M.P. (1997): "Altered neural cell fates and medullablastom in mouse patched mutants", *Science* **277**, 1109-1113). This cascade of signalling events, best characterized in *Drosophila*, also involves a number of intracellular components including *fused* (a serine threonine kinase), *suppressor of fused*, *costal 2*, and *cubitus interruptus* (Ruiz i Altaba, A., "Catching a Gli-mpse of Hedgehog" (1997) *Cell* **90**, 193-196). The latter is a transcription factor that positively regulates the expression of target genes which also include PTCH itself.

Mutations in the PTCH gene have been identified in both sporadic and familial BCCs (Gailani, M.R., Stähle-Bäckdahl, M., Leffell, D.J., Glynn, M., Zaphiropoulos, P.G., Pressman, C., Undén, A.B., Dean, M., Brash, D. E., Bale, A.E. and Toftgård, R. (1996): "The role of human homologue of *Drosophila* patched in sporadic basal cell carcinomas" *Nature Genet.* **14**, 78-81). The lack of the normal PTCH protein in these cells allows the constitutive signalling of SMO to occur, resulting in the accumulation of mutant PTCH mRNAs (Undén, B. A., Zaphiropoulos, P.G., Bruce, K., Toftgård, R., and Stähle-Bäckdahl, M. (1997): "Human patched (PTCH) mRNA is overexpressed consistently in tumor cells of both familial and sporadic basal cell carcinoma", *Cancer Res.* **57**, 2336-2340).

WO 96/11260 discloses the isolation of *patched* genes and the use of the PTC protein to identify ligands, other than the established ligand Hedgehog, that bind thereto.

5 However, there is still a need of a further understanding of the SHH/PTCH cell signalling, which may be provided by disclosure of further genes, peptides and proteins involved therein.

Summary of the invention

10 The present invention provides a significant step forward regarding the understanding of the above described pathway. By a combination of cDNA library and RACE analysis a novel human *patched*-like gene (PTCH2) has been cloned and sequenced. Several alternatively spliced mRNA forms of PTCH2 have been identified, including transcripts lacking segments thought to be involved in sonic hedgehog
15 (SHH) binding and mRNAs with differentially defined 3' terminal exons. Accordingly, the invention relates to isolated such mRNAs as well as to cDNAs complementary thereto.

Brief description of the drawings

20 Figure 1 shows the genomic sequence of SEQ ID NO:5, wherein exons and introns are designated in the genomic sequence of the novel human patched 2 gene.

Figure 2A discloses an amino acid sequence comparison of the human PTCH2 (upper lines) and PTCH1 (lower lines) sequences.

Figure 2B is a representation of the alternative splicing events that result in different C-termini.
25

Figure 2C is a representation of the different variations of spliced transcripts encompassing exon 1 and exon 2 sequences.

Figure 3A is a dark-field photomicrograph of a BCC tumor hybridised with ³⁵S-labeled antisense probe showing abundant signal for PTCH1 mRNA (light grains) in
30 all BCC tumor cells.

Figure 3B discloses PTCH2 mRNA overexpression in BCC and is in contrast mainly expressed in the basaloid cells in the periphery of the tumor nests.

Figure 3C is another BCC showing a strong PTCH2 mRNA signal in the periphery of the tumor nest (Tu), whereas no signal is detected in epidermis (Ep).

5 Figure 3D are sections of the same tumor (C) hybridised with the PTCH2 sense probe showed no signal.

Figure 3E shows immunoreactivity for Ki-67.

Figure 3F discloses how tumor nests under high power magnification demonstrate abundant PTCH2 mRNA signal (black grains) in the dark basaloid tumor cells and
10 lower signal in the center (arrow).

Definitions

The terms "polypeptide", "peptide" and "protein" are used interchangeably herein to refer to a polymer of amino acid residues. The terms apply to amino acid polymers in which one or more amino acid residue is an artificial chemical analogue of a
15 corresponding naturally occurring amino acid, as well as to naturally occurring amino acid polymers.

The terms "isolated" "purified" or "biologically pure" refer to material which is
20 substantially or essentially free from components which normally accompany it as found in its native state.

The term "nucleic acid" refers to a deoxyribonucleotide or ribonucleotide polymer in either single- or double-stranded form, and unless otherwise limited, encompasses known analogs of natural nucleotides that can function in a similar manner as
25 naturally occurring nucleotides.

A "label" is a composition detectable by spectroscopic, photochemical, biochemical, immunochemical, or chemical means. For example, useful labels include ^{32}P ,
30 fluorescent dyes, electron-dense reagents, enzymes (e.g., as commonly used in a ELISA), biotin, dioxigenin, or haptens and proteins for which antisera or mono-

clonal antibodies are available (*e.g.*, the peptide of SEQ ID NO:1 can be made detectable, *e.g.*, by incorporating a radio-label into the peptide, and used to detect antibodies specifically reactive with the peptide).

5 As used herein a "nucleic acid probe" is defined as a nucleic acid capable of binding to a target nucleic acid of complementary sequence through one or more types of chemical bonds, usually through complementary base pairing, usually through hydrogen bond formation. As used herein, a probe may include natural (*i.e.* A, G, C, or T) or modified bases (7-deazaguanosine, inosine, *etc.*) In addition, the bases in a
10 probe may be joined by a linkage other than a phosphodiester bond, so long as it does not interfere with hybridisation. Thus, for example, probes may be peptide nucleic acids in which the constituent bases are joined by peptide bonds rather than phosphodiester linkages. It will be understood by one of skill in the art that probes may bind target sequences lacking complete complementarity with the probe sequence depending upon the stringency of the hybridisation conditions. The probes
15 are preferably directly labeled as with isotopes, chromophores, lumiphore, chromogens, or indirectly labeled such as with biotin to which a streptavidin complex may later bind. By assaying for the presence or absence of the probe, one can detect the presence or absence of the selct sequence or subsequence.

20

A "labeled nucleic acid probe" is a nucleic acid probe that is bound, either covalently, through a linker, or through ionic, van der Waals or hydrogen bonds to a label such that the presence of the probe may be detected by detecting the presence of the label bound to the probe.

25

The term "target nucleic acid" refers to a nucleic acid (often derived from a biological sample), to which a nucelic acid probe is designed to specifically hybridise. It is either the presence or absence of the target nucleic acid that is to be detected, or the amount of the target nucleic acid that is to be quantified. The target nucleic acid has
30 a sequence that is complementary to the nucleic acid sequence of the corresponding probe directed to the target. The term target nucleic acid may refer to the specific

subsequence of a larger nucleic acid to which the probe is directed or to the overall sequence (*e.g.*, gene or mRNA) whose expression level it is desired to detect. The difference in usage will be apparent from context.

- 5 The term "recombinant" when used with reference to a cell, or nucleic acid, or vector, indicates that the cell, or nucleic acid, or vector, has been modified by the introduction of a heterologous nucleic acid or the alteration of a native nucleic acid, or that the cell is derived from a cell so modified.
- 10 The term "identical" in the context of two nucleic acids or polypeptide sequences refers to the residues in the two sequences which are the same when aligned for maximum correspondence. Optimal alignment of sequences for comparison can be conducted, *e.g.*, by the local homology algorithm of Smith and Waterman (1981) *Adv. Appl. Math.* 2: 482, by the homology alignment algorithm of Needleman and
- 15 Wunsch (1970) *J. Mol. Biol.* 48:443, by the search for similarity method of Pearson and Lipman (1988) *Proc. Natl. Acad. Sci. USA* 85: 2444, by computerized implementations of these algorithms (GAP, GESTFIT, FASTA, and TFASTA in the Wisconsin Genetics Software Package, Genetics Computer Group, 575 Science Dr., Madison, WI) or by inspection. The BLAST algorithm performs a statistical analysis of the similarity between two sequences; *see e.g.*, Karlin and Altschul (1993) *Proc. Nat'l Acad. Sci. USA* 90: 5873-5787.
- 20

- The term "substantial identity" or "substantial similarity" in the context of a polypeptide indicates that a polypeptides comprises a sequence with at least 70% sequence identity to a reference sequence, or preferably 80%, or more preferably 85% sequence identity to the reference sequence, or most preferably 90% identity over a comparison window of about 10-20 amino acid residues. An indication that two polypeptide sequences are substantially identical is that one peptide is immunologically reactive with antibodies raised against the second peptide. Thus, a polypeptide is substantially identical to a second polypeptide, for example, where the two peptides differ only by a conservative substitution.
- 25
- 30

An indication that two nucleic acid sequences are substantially identical is that the polypeptide which the first nucleic acid encodes is immunologically cross reactive with the polypeptide encoded by the second nucleic acid. Another indication that two nucleic acid sequences are substantially identical is that the two molecules hybridise to each other under stringent conditions.

The phrase "hybridising specifically to", refers to the binding, duplexing, or hybridising of a molecule only to a particular nucleotide sequence under stringent conditions when that sequence is present in a complex mixture (*e.g.*, total cellular DNA or RNA). The term "stringent conditions" refers to conditions under which a probe will hybridise to its target subsequence, but to no other sequences. Stringent conditions are sequence-dependent and will be different in different circumstances. Longer sequences hybridise specifically at higher temperatures. Generally, stringent conditions are selected to be about 5°C lower than the thermal melting point T_m for the specific sequence at a defined ionic strength and pH. The T_m is the temperature (under defined ionic strength, pH, and nucleic acid concentration) at which 50% of the probes complementary to the target sequence hybridise to the target sequence at equilibrium. (As the target sequences are generally present in excess, at T_m , 50% of the probes are occupied at equilibrium). Typically, stringent conditions will be those in which the salt concentration is less than about 1.0 M Na ion, typically about 0.01 to 1.0 M Na ion concentration (or other salts) at pH 7.0 to 8.3 and the temperature is at least about 30°C for short probes (*e.g.*, 10 to 50 nucleotides) and at least about 60°C for long probes (*e.g.*, greater than 50 nucleotides). Stringent conditions may also be achieved with the addition of destabilizing agents such as formamide.

The term "antibody" refers to a polypeptide substantially encoded by an immunoglobulin gene or immunoglobulin genes, or fragments thereof which specifically bind and recognize an analyte (antigen).

A "chimeric antibody" is an antibody molecule in which (a) the constant region, or a portion thereof, is altered, replaced or exchanged so that the antigen binding site (variable region) is linked to a constant region of a different or altered class, effector function and/or species, or an entirely different molecule which confers new properties to the chimeric antibody, *e.g.*, an enzyme, toxin, hormone, growth factor, drug, *etc.*; or (b) the variable region, or a portion thereof, is altered, replaced or exchanged with a variable region having a different or altered antigen specificity.

The term "immunoassay" is an assay that utilizes an antibody to specifically bind an analyte. The immunoassay is characterized by the use of specific binding properties of a particular antibody to isolate, target, and/or quantify the analyte.

The phrases "specifically binds to a protein" or "specifically immunoreactive with", when referring to an antibody refers to a binding reaction which is determinative of the presence of the protein in the presence of a heterogeneous population of proteins and other biologics. Thus, under designated immunoassay conditions, the specified antibodies bind preferentially to a particular protein and do not bind in a significant amount to other proteins present in the sample. Specific binding to a protein under such conditions requires an antibody that is selected for its specificity for a particular protein. A variety of immunoassay formats may be used to select antibodies specifically immunoreactive with a particular protein. For example, solid-phase ELISA immunoassays are routinely used to select monoclonal antibodies specifically immunoreactive with a protein. *See harlow and Lane (1988) Antibodies, A Laboratory Manual*, Cold Spring Harbour Publications, New York, for a description of immunoassay formats and conditions that can be used to determine specific immunoreactivity.

A "gene product", as used herein, refers to a nucleic acid whose presence, absence, quantity, or nucleic acid sequence is indicative of a presence, absence, quantity, or nucleic acid composition of the gene. Gene products thus include, but are not limited to, and mRNA transcript and cDNA reverse transcribed from an mRNA, and RNA

transcribed from that cDNA, a DNA amplified from the cDNA, an RNA transcribed from the amplified DNA or subsequences of any of these nucleic acids. Polypeptides expressed by the gene or subsequences thereof are also gene products. The particular type of gene product will be evident from the context of the usage of the term.

A "modified drug" means a compound, which retains the pharmaceutical properties of the original drug or active substance while the structure thereof has been modified. Further, encompassed by the term "drug" are also compounds useful in diagnostic methods by their specific binding properties.

Detailed description of the invention

In a first aspect, the present invention relates to an isolated human protein, or an analogue or a variant thereof, capable of participating in the human PTCH/SHH pathway during embryonic development and/or carcinogenesis, such as basal cell carcinoma. The novel protein according to the invention is encoded by a novel gene, which isolated nucleic acid is described in detail below and which is denoted *patched 2* (PTCH2) due to its similarities with *patched 1* (PTCH1). Accordingly, the protein according to the invention exhibits substantial differences in sequence and functions when compared to human PTCH1 protein. The protein according to the invention is best characterized by its functions which when compared to human PTCH1 are similar but distinct therefrom in certain ways, more specifically disclosed below in the section "Results and discussion". The novel human PTCH2 protein according to the invention is also distinct from the previously isolated mouse PTCH2. Thus, in the preferred embodiment thereof, it comprises a substantial part of the amino acid sequence disclosed in SEQ ID NO: 1 and submitted to the GenBank under protein id no AAD17260.1. even though it is to be understood that the present invention encompasses any fragment, analogue or variant thereof exhibiting the biological functions of the PTCH2 protein disclosed herein. Thus, preferably, the present protein comprises at least about 1000, more preferably at least about

1040 and most preferably essentially all of the amino acids of the sequence denoted SEQ ID NO: 1, such as about 1100.

5 The proteins according to the invention are easily prepared by someone skilled in this field by recombinant DNA techniques using the molecules disclosed below or any synthetic method (*see e.g.* Barany and Merrifield, Solid-Phase Peptide synthesis, pp. 3-284 in *The Peptides: Analysis, Synthesis, Biology*, Vol. 2: Special Methods in Peptide synthesis, Part A, Merrifield *et al.*, J. Am. Chem. Soc., 2149-2156).

10 The present invention also relates to the use of the peptides, polypeptides and proteins disclosed herein as lead compounds in methods aimed at finding novel substances, i.e. modified drugs, such as substances exhibiting equivalent or even more advantageous properties than the lead compounds as such. Such modified drugs may
15 also be designed by methods of combinatorial chemistry, wherein a structurally similar compound is specifically designed e.g. by aid of computers. Alternatively, the present modified drug is identified by screening of a library of candidate compounds, e.g. using an antibody according to the invention. In the present context, it is to be understood that when such a modified drug has been identified, it is possible
20 to produce it by any other suitable technique. The invention also relates to proteomic methods wherein the present molecules are used as well as to such a use *per se*.

A second aspect of the present invention is a nucleic acid encoding a protein, an analogue or a variant thereof as defined above, that is, the protein coding region of
25 the novel human isolated PTCH2 gene. The PTCH2 gene is 57% identical to PTCH1 and 91% identical to the published mouse *Ptch2* sequence (*see* Motoyama *et al.*, (1998), *supra*). Thus, preferably, the nucleic acid according to the present invention comprises at least about 3000 bases, more preferably at least about 3094 bases and most preferably essentially all of the sequence denoted SEQ ID NO: 2.

In a specific aspect, the present invention relates to the isolated human genomic PTCH2 nucleic acid comprising parts or all of the genomic sequence denoted SEQ ID NO: 5. In the disclosure of the genomic sequence shown in Fig 1, the exon/intron structure of the present gene is shown. Further to the exons shown therein, exon 12a and 12b has also been identified, as specifically defined by SEQ ID NO:3 and SEQ ID NO:4, respectively. Interestingly, there is a splice variant that joins exon 12a to a 3' segment of exon 12b with conservation of the intronic GT-AG dinucleotides. Exons 12a and 12b are not variants, but the actual exons of the gene identified by sequencing the corresponding genomic region. (Materials and methods were as discribed below). Accordingly, these findings show that PTCH2 has the same intron/exon structure organization as PTCH1. In another embodiment of this aspect, the present invention relates to a transcript that has skipped only one of the exons 9 and 10 defined in Fig 1. In an alternative embodiment, the transcript according to the invention has skipped both of exon 9 and 10. The splice variants of the present gene are discussed in more detail below in the section "Results", all of which are included within the scope of the present invention. This aspect of the invention advantageously enables design of suitable PCR primers, which in turn enables screening for mutations of all of the coding sections thereof, *e.g.* by SSCP analysis, sequencing, or any other suitable method known to someone skilled in this field. Thus, the novel human PTCH2 gene according to the invention has been localized by radiation hybrid mapping to chromosome 1p32-35 with D1S211 and WI-1404 as closest flanking markers and with an estimated localization 5.5cR from D1S443. This region is often lost by LOH in various different tumor types, such as neuroblastoma, melanoma, breast cancer, colon cancer *etc.* Accordingly, PTCH2 is a candidate for a tumor suppressor gene in this region and the present invention also encompass diagnostic methods based on this new disclosure.

To this chromosomal region, three cancer predisposition syndromes have also been mapped, namely, familial melanoma CMM1, modifier locus for familial adenomatous polyposis hMom1 and Michelin Tire Baby Syndrome. PTCH2 is further a candidate for the gene behind these heritary syndromes. The present molecules are the-

referred advantageously used in the context of these conditions, *e.g.* in therapy and/or diagnosis, such as in assays.

Further, the invention also relates to various PCR primers based on intronic sequences, allowing amplification of all coding sequence. Such primers are advantageously used for mutation screening.

Further, the present invention also relates to the any isolated nucleic acid capable of specifically hybridising to a nucleic acid according to the invention. In addition, the invention also relates to such an isolated nucleic acid which comprises one or more mutations compared to the genomic sequence as well as the use of the novel isolated nucleic acids, *e.g.* to identify mutations for diagnostic and/or therapeutic purposes.

Further embodiments of this aspect of the invention includes nucleic acid probes, *e.g.* DNA probes, labelled nucleic acids, cDNAs, RNAs *etc.*, that is, all gene products obtainable by someone skilled in this field based on the novel isolated human PTCH2 gene.

Another aspect of the invention is a nucleic acid corresponding to any one of the splicing variants disclosed in Figure 2B, a protein or polypeptide encoded thereof as well as various uses thereof.

As regards the preparation of nucleic acids according to the invention, any suitable recombinant DNA technique or synthetic method may be used. (For general laboratory procedures useful in this context, *see e.g.* Sambrook *et al.*, Molecular Cloning, A Laboratory Manual, 2nd ed., Vol. 1-3, Cold Spring Harbor Laboratory, Cold Spring Harbor, NY, 1989; Berger and Kimmel, Guide to Molecular Cloning Techniques, Methods in Enzymology, Vol. 153, Academic Press, Inc., San Diego, CA; Current Protocols in Molecular Biology, F.M. Ausbel *et al.*, eds., Current Protocols (1994)).

A further aspect of the present invention is a vector comprising a nucleic acid as defined above. Vectors are *e.g.* useful for transforming cells in vitro or in vivo to express the proteins and peptides according to the invention and may *e.g.* be plasmids, viruses *etc.*

5

Another aspect of the invention is a recombinant cell, such as a eucaryotic, *e.g.* a mammalian cell, or a procaryotic cell, *e.g.* a bacteria, comprising a vector as defined above. Such cells may *e.g.* be used to monitor expression levels of the proteins and polypeptides according to the invention in a wide variety of contexts. For example, when the effects of a drug is to be determined, the drug will be administered to the transformed organism, tissue or cell. Accordingly, model systems including such cells are another aspect of the invention.

10

15

A further aspect of the invention is an antibody, such as a monoclonal or polyclonal antibody, which specifically binds to a protein or polypeptide according to the invention. An exemplary immunoglobulin (antibody) structural unit comprises a tetramer. Each tetramer is composed of two identical pairs of polypeptide chains, each pair having one "light" (about 25 kD) and one "heavy" chain (about 50-70 kD). The N-terminus of each chain defines a variable region of about 100 to 110 or more amino acids primarily responsible for antigen recognition. The terms variable light chain (V_L) and variable haecavy chain (V_H) refer to these light and heavy chains, respectively.

20

25

The invention also encompasses chimeric or other antibodies that binds the present proteins or polypeptides. Further, the invention also relates to the use of the present antibodies in assays. (In this context, *see e.g. Fundamental Immunology*, Third Edition, W.E. Paul, ed., Raven Press, N.Y. 1993).

30

Further, the invention also relates to a recombinant cell expressing an antibody according to the invention.

In general, prokaryotes can be used for cloning the DNA sequences encoding a human anti-PTCH2 immunoglobulin chain. *E. coli* is one prokaryotic host particularly useful for cloning the DNA sequences of the present invention. Microbes, such as yeast are also useful for expression. *Saccharomyces* is a preferred yeast host, with
5 suitable vectors having expression control sequences, an origin of replication, termination sequences and the like as desired. Typical promoters include 3-phosphoglycerate kinase and other glycolytic enzymes. Inducible yeast promoters include, among others, promoters from alcohol dehydrogenase 2, isocytochrome C, and enzymes responsible for maltose and galactose utilization.

10 Mammalian cells are a particularly preferred host for expressing nucleotide segments encoding immunoglobulins or fragments thereof (*see, e.g.* Winnacker, *From Genes to Clones*, VCH Publishers, N.Y., 1987). A number of suitable host cell lines capable of secreting intact heterologous proteins have been developed in the art, and
15 include CHO cell lines, various COS cell lines, HeLa cells, L cells and myeloma cell lines. Preferably, the cells are nonhuman. Expression vectors for these cells can include expression control sequences, such as an origin of replication, a promoter, an enhancer (Queen *et al.* (1986) *Immunol. Rev.* 89:49), and necessary processing information sites, such as ribosome binding sites, RNA splice sites, polyadenylation
20 sites, and transcriptional terminator sequences. Preferred expression control sequences are promoters derived from endogenous genes, cytomegalovirus, SV40, adenovirus, bovine papillomavirus, and the like (*see, e.g.*, Co *et al.* (1992) *J. Immunol.* 1458: 1149).

25 An additional aspect of the present invention is a kit for the detection of a human PTCH2 gene or polypeptide comprising in a container a molecule selected from the group consisting of a nucleic acid, a polypeptide or a protein or an antibody according to the invention. Further suitable components of such a kit are easily determined by someone skilled in this field as are the conditions for the use thereof.

Further, the invention also relates to the use of a nucleic acid selected from the group consisting of SEQ ID NOS: 2-4 and SEQ ID NO: 5 in gene therapy. In addition to said specifically disclosed sequences, any one of the herein disclosed exons may be used to this end. For a review of gene therapy procedures, see Anderson, 5 *Science* (1992) 256:808-813; Nabel and Felgner (1993) *TIBTECH* 11: 211-217; Mitani and Caskey (1993) *TIBTECH* 11: 162-166; Mulligan (1993) *Science* 926-932; Dillon (1993) *TIBTECH* 11: 167-175; Miller (1992) *Nature* 357: 455-460; Van Brunt (1988) *Biotechnology* 6(10): 1149-1154; Vigne (1995) *Restorative Neurology and Neuroscience* 8: 35-36; Kremer and Perricaudet (1995) *British Medical Bulletin* 10 51(1) 31-44; Haddada *et al.* (1995) in *Current Topics in Microbiology and Immunology* Doerfler and Böhm (eds) Springer-Verlag, Heidelberg Germany; and Yu *et al.*, *Gene Therapy* (1994) 1:13-26.

Delivery of the gene or genetic material into the cell is the first critical step in gene 15 therapy treatment of disease. A large number of delivery methods are well known to those of skill in the art. Such methods include, for example liposome-based gene delivery (Debs and Zhu (1993) WO 93/24640; Mannino and Gould-Fogerite (1988) *BioTechniques* 6(7): 682-691; Rose U.S. Pat No. 5,279,833; Brigham (1991) WO 91/06309; and Felgner *et al.* (1987) *Proc. Natl. Acad. Sci. USA* 84: 7413-7414), and 20 replication-defective retroviral vectors harboring a therapeutic polynucleotide uence as part of the retroviral genome (see, e.g., Miller *et al.* (1990) *Mol. Cell. Biol.* 10:4239 (1990; Kolberg (1992) *J. NIH Res.* 4:43, and Cornetta *et al.* *Hum. Gene Ther.* 2:215 (1991)). Widely used retroviral vectors include those based upon murine leukemia virus (MuLV), gibbon ape leukemia virus (GaLV), Simian Immuno deficiency virus (SIV), human immuno deficiency virus (HIV), and combinations 25 thereof. See, e.g., Buchscher *et al.* (1992) *J. Virol.* 66 (5) 2731-2739; Johann *et al.* (1992) *J. Virol.* 66 (5):1635-1640 (1992); Sommerfelt *et al.*, (1990) *Virol.* 176:58-59; Wilson *et al.* (1989) *J. Virol.* 63:2374-2378; Miller *et al.*, *J Virol.* 65:2220-2224 (1991); Wong-Staal *et al.*, PCT/US94/05700, and Rosenberg and Fauci (1993) 30 in *Fundamental Immunology, Third Edition* Paul (ed) Raven Press, Ltd., New York and the references therein, and Yu *et al.*, *Gene Therapy* (1994) supra).

The present invention may also be used in the pharmaceutical industry. For example, it will provide information that eventually may enable cells from fetal tissue, which may be transplanted into patients suffering from *e.g.* Parkinson's disease or cancer, such as BCC. (For a brief review of methods of drug delivery, *see* Langer
5 249:1 527-1533 (1990), Remington's Pharmaceutical Sciences, Mack Publishing Company, Philadelphia, PA, 17th ed. (1985) *etc.*)

Detailed description of the drawings

Figure 1 shows the genomic sequence of SEQ ID NO:5, wherein exons and introns
10 are designated in the genomic sequence of the present human patched 2 gene. However, exons 12a and 12b discussed above are not specifically shown in Figure 1, but is instead disclosed as the separate sequences SEQ ID NO:3 and SEQ ID NO:4, respectively. Figure 2A discloses an amino acid sequence comparison of the human PTCH2 (upper lines) and PTCH1 (lower lines) sequences. Vertical lines indicate
15 identical amino acids, while dots similar amino acids. The PTCH2 sequence presented is composed of the original cDNA clones and of the products of the 5' RACE analysis.

Figure 2B is a representation of the alternative splicing events that result in different C-termini. In the parotid gland and the colon, the penultimate and the last exon
20 are canonically joined together. In fetal brain however the penultimate exon with part of the 3' intron functions as the terminal exon. The intronic sequence is shown by small letters with the flanking exonic by capital letters. Above the nucleotide sequence, the deduced amino acid sequence is shown, and below is the corresponding
25 sequence of the mouse Ptch2. The conserved intronic dinucleotides are shown by bold letters and the termination signals are indicated by asterisks. Note the absence of conservation of the position of the termination codons between the mouse and human PTCH2 sequences. The putative polyadenylation signals are also shown in this diagram. The genomic organization was obtained by analyzing BAC clones en-
30 compassing the PTCH2 gene.

Figure 2C is a representation of the different variations of spliced transcripts encompassing exon 1 and exon 2 sequences. The canonical exons 1 and 2 are shown by boxes and the intron between them by a solid line. The GT and AG dinucleotides spanning the sequences that are used as introns in individual transcripts are indicated by small letters. G, Genomic structure, derived from sequencing segments of BAC clones encompassing the PTCH2 gene; C, Canonical transcript; A, Transcript A (the skipped exons 9 and 10 of this product are not shown in the diagram); B, Transcript B.

10 Figure 3A is a dark-field photomicrograph of a BCC tumor hybridised with ³⁵S-labeled antisense probe showing abundant signal for PTCH1 mRNA (light grains) in all BCC tumor cells.

Figure 3B discloses PTCH2 mRNA overexpression in BCC and is in contrast mainly expressed in the basaloid cells in the periphery of the tumor nests.

Figure 3C is another BCC showing a strong PTCH2 mRNA signal in the periphery of the tumor nest (Tu), whereas no signal is detected in epidermis (Ep).

20 Figure 3D are sections of the same tumor (C) hybridised with the PTCH2 sense probe showed no signal.

Figure 3E shows immunoreactivity for Ki-67 (brown precipitate) seen in the periphery, in the cells that showed strong upregulation of PTCH2 mRNA.

25 Figure 3F discloses tumor nests under high power magnification demonstrate abundant PTCH2 mRNA signal (black grains) in the dark basaloid tumor cells and lower signal in the center (arrow). Bars (A-E), 24 µm, and F, 6 µm.

EXPERIMENTAL

Materials and methods

In the present context, a general reference is made to G. Zaphiropoulos et al., Cancer Res., vol. 59, p. 787-792, February 15, 1999, disclosing useful methods in the present context. All references mentioned in the present application are hereby included herein by reference. The examples below are not intended to limit the scope of the invention but merely as an illustration.

The RACE analysis was performed essentially as described before (Zaphiropoulos, P.G. and Toftgård, R. (1996): "cDNA cloning of a novel WD repeat protein mapping to the 9q22.3 chromosomal region", DNA Cell Biol. 15, 1049-1056) using the Marathon kit (Promega). The primer sequences used for RACE are available upon request.

The PTCH2, 35S-labeled RNA probes used for the in situ hybridisations, that were performed as previously described (Undén *et al.*, (1997), *supra*), corresponded to positions 218 to 437 and 838 to 920 in the PTCH2 sequence of SEQ ID NO:1 .

Results and discussion

In order to identify additional components of the PTCH/SHH cascade of signalling events, the Incyte LifeSeqTM database (Incyte Pharmaceuticals Inc., Palo Alto, CA, USA) was searched using PTCH sequences. In addition to clones representing the PTCH cDNA, two nearly identical cDNAs were identified, from the parotid gland and the colon, that contained sequences similar to, but distinct from, the 3' end of PTCH. By 5' RACE analysis using fetal brain cDNAs additional sequence information from these transcripts (termed PTCH2) and corresponding to a full length cDNA, was obtained (Fig. 2A). PTCH2 is 57% identical to PTCH1, with a significantly variable region present between the transmembrane domains 6 and 7, and 91% identical to the recently published mouse Ptch2 sequence (Motoyama, J., Takabatake, T., Takeshima, K. and Hui, C. (1998): "Ptch2, a second mouse Patched gene is co-expressed with Sonic hedgehog", Nature Genet. 18, 104-106). In simila-

rity with the mouse gene, PTCH2 lacks the C-terminal extension present in human, mouse and chicken PTCH1 (Goodrich, L.V., Johnson, R.L., Milenkovic, L., McMahon, J.A., and Scott, M.P. (1996): "Conservation of the hedgehog/patched signalling pathway from flies to mice: Induction of a mouse patched gene by Hedgehog", *Genes Dev.* 10, 301-312, Marigo, V., Scott, M.P., Johnson, R.L., Goodrich, L.V. and Tabin, C.J. (1996): "Conservation in hedgehog signalling: Induction of a chicken patched homolog by Sonic hedgehog in the developing limb", *Development* 122, 1225-1233). However, according to the present invention, it has been shown that the human PTCH2 cDNA terminates 36 amino acids earlier than the mouse Ptc2 sequence. Moreover, when 3' RACE was performed from fetal brain, an alternate C-terminal region was identified. This had a high structural similarity with the mouse Ptc2 C-terminal sequence and originates from the genomic region that links the last two exons of PTCH2 (Fig. 2B). Therefore, in these alternatively spliced transcripts, the penultimate exon with a segment of the contiguous 3' intron serves as the terminal exon.

Moreover the human and mouse transcripts differed in the position of the termination signals (the human sequence is 21 amino acids longer), suggesting a non-conserved, species-specific function of this alternate C-terminal domain. The finding of two possible C-terminal regions for PTCH2 is intriguing and implies a role of this phenomenon in modulating signalling. Additional alternatively spliced transcripts were also identified by the RACE analysis (Fig. 2C). Transcript A lacks the sequence that corresponds to exons 9 and 10 of PTCH1 (preliminary comparisons of the intron-exon junctions of PTCH2 with PTCH1 indicate a similar genomic organization), with the open reading frame being retained at the exon 8 to exon 11 junction. Exons 9 and 10 code for the last part of the first extracellular loop and for transmembrane domains 2 and 3 in the putative structure of the PTCH1 protein. Furthermore this transcript also lacks a 5' segment of the canonical exon 2, due to the use of an alternative 3' splice site present in this exon, with the open reading frame being maintained. The functional consequence of this alternative splicing is not yet known, but it is interesting to note that the extracellular loops in PTCH1 are

presumed to be involved in binding of the ligand SHH (Marigo *et al.*, (1996), Nature 384, supra; Stone *et al.*, (1996), Nature 384, supra) and that insertion of a neo-cassette in intron 9 of the mouse PTCH1 gene is associated with a severe phenotype (Hahn, H., Wojnowski, L., Zimmer, A.M., Hall, J., Miller, G. and Zimmer, A. (1998): "Rhabdomyosarcomas and radiation hypersensitivity in a mouse model of Gorlin syndrome", Nature Med. 4, 619-622). Furthermore, exons 9 and 10 encode part of a putative sterol sensing domain (Osborne, T.F. and Rosenfeld, J.M. (1998): "Related membrane domains in proteins of sterol sensing and cell signalling provide a glimpse of treasures still buried within the dynamic realm of intracellular metabolic regulation", Curr. Opin. Lipidol. 9, 137-140, also found in PTCH1, and which has recently been implicated in mediating the potent modulating effect of cholesterol on SHH/PTCH signalling (Cooper, M.K., Porter, J.A., Young, K.E., and Beachy, P.A. (1998): "Teratogen-mediated inhibition of target tissue response to Shh signalling", Science 280, 1603-1607). Thus, if PTCH2 also serves as a receptor for SHH and/or related factors, the receptor form lacking exons 9 and 10 may show altered signalling properties. Transcript B contains additional sequences between canonical exons 1 and 2, that originate from the 5' end of intron 1. The open reading frame that includes the initiator methionine of exon 1 is not maintained in this transcript, suggesting that, if this transcript is functional, either the methionine in exon 2 or non-methionine codons are used in order to produce a protein product, in similarity to what has been proposed for the alternative spliced products of human PTCH1 (Hahn *et al.*, Cell 85, supra). By radiation hybrid mapping the PTCH2 gene was localized to the short arm of chromosome 1, in difference to PTCH1 residing on chromosome 9q22.3.

The mouse and zebrafish homologs of PTCH2 have been reported to be expressed in a partly overlapping pattern with PTCH1 during embryonic development and to be induced by SHH (Motoyama *et al.*, (1998) Nature Genet. 18, supra, Concordet, J.P., Lewis, K.E., Moore, J.W., Goodrich, L.V., Johnson, R.L., Scott, M.P., and Ingham, P.W. (1996): "Spatial regulation of a zebrafish patched homologue reflects

the roles of sonic hedgehog and protein kinase A in a neural tube and somite patterning", Development 122, 2835-2846), implicating a role in this signalling pathway. We were with this background interested to analyze the expression of PTCH2 in BCCs which show consistent upregulation of PTCH1 in all tumor cells (Undén *et al.*, (1997) Cancer res. 57, supra). In situ hybridisation was performed on six familial and four sporadic BCCs of different histological subtypes. A strong positive signal for PTCH2 mRNA was observed exclusively in the tumor cells of all BCCs. Notably, the signal was consistently stronger in the palisading peripheral cells of the tumor nests (Fig. 2). These cells also showed a positive immunostaining for the cell proliferation marker, Ki-67.

The finding that in BCCs having frequent mutations in the PTCH1 gene, the expression of the PTCH2 mRNAs is upregulated, tightly links the novel PTCH2 according to the invention with the PTCH/SHH cascade of signalling events. It is therefore likely that PTCH2 represents a target gene of this pathway which is under the negative regulation of PTCH1, precisely as PTCH1 itself. Moreover this observation strongly suggests that PTCH2 has functions distinct from PTCH1 since upregulation of PTCH2 expression appears unable to compensate for inactive PTCH1 protein. This conclusion is also supported by the early embryonic lethality seen in PTCH1 (-/-) mice 5,13) and the lack of genetic heterogeneity in Gorlin syndrome. However, whether PTCH2 may block the constitutive signalling of SMO, or could act as an additional SHH receptor, possible dependent on alternative splicing, remains as the subject of further experimentation.

CLAIMS

1. An isolated human protein or an analogue or variant thereof capable of participating in the human PTCH/SHH pathway during embryonic development and/or carcinogenesis comprising at least about 1040 amino acids of SEQ ID NO: 1.
- 5 2. A protein according to claim 1, which is essentially comprised of SEQ ID NO: 1.
3. A nucleic acid encoding a protein according to any one of claims 1 and 2.
4. A nucleic acid encoding a protein according to claim 3 comprising at least about 3094 bases of SEQ ID NO: 2.
- 10 5. An isolated genomic nucleic acid comprising parts or all of SEQ ID NO: 5.
6. An isolated variant of the nucleic acid according to claim 5.
7. A nucleic acid having the sequence of a splicing variant selected from the group of variants defined in Figure 2B and SEQ ID NO:3 and SEQ ID NO:4.
8. An isolated nucleic acid capable of specifically hybridising to a nucleic acid according to any one of claims 3-7.
- 15 9. A protein or polypeptide encoded by a nucleic acid according to claim 7.
10. A protein according to claim 1 or 2, a nucleic acid according to any one of claims 2-8 or or a polypeptide according to claim 9, for use as a medicament.
11. Use of a protein according to claim 1 or 2, a nucleic acid according to any one of claims 2-8 or a polypeptide according to claim 9 in the manufacture of a medicament for the treatment of a condition involving tumors, such as BSS.
- 20 12. A method of in vitro or in vivo diagnosis, wherein a protein according to claim 1 or 2, a nucleic acid according to any one of claims 2-8 or a polypeptide according to claim 9 is used.
- 25 13. A method of screening wherein a library of suitable candidate compounds is screened for modified drugs using a protein according to claim 1 or 2 or a polypeptide according to claim 9 as a lead compound.
14. A method of synthesis of a modified drug, wherein a protein according to claim 1 or 2, or a protein or polypeptide according to claim 9, is used.
- 30 15. A modified drug identified by the method according to claim 13 or synthesized according to claim 14.

16. A vector comprising a nucleic acid according to any one of claim 3-8.
17. A recombinant cell comprising a vector according to claim 16.
18. An antibody which specifically binds to a protein or polypeptide according to claim 1, 2 or 9.
- 5 19. A recombinant cell expressing an antibody according to claim 18.
20. A kit for the detection of a human PTCH2 gene or polypeptide comprising in a container a molecule selected from the group consisting of a nucleic acid according to any one of claims 3-8, a polypeptide or protein according to claim 1, 2 or 9 or an antibody according to claim 18.
- 10 21. Use of a nucleic acid according to any one of claims 3-7 in gene therapy.
22. Use of nucleic acid according to any one of claims 3-7 as a probe, a primer or a diagnostic reagent.

The intron sequences between exons 2 - 3 and exons 18 - 19 are missing (introns: small letters, exons: capital letters). Small letters in the first exon indicate nucleotides that have not been unambiguously determined.

Exon 1

1 CGGGTGAATC CCGGCGCCGC GCCCGGACC CGCAGCTCCC TGCACCTCTC
51 CCTCCCAGCC GCTTTAACAC CCACACCCCA CAGTCTCTCC CACGsCCGCG
101 CCTTGGCGGC CCCACTGAAT CCCTACGCGG GGCCCAGCGG TACCGGGAGA
151 CCGGGCTAGC CTATGGGAGC GCCCAGATAA CGCGGGTTGG GGGCGCCCCG
201 GCCCcCATCC CCGCCAGCA**T G**ACTCGATCG CCGCCCCCTCA GAGAGCTGCC
251 CCCGAGTTAC ACACCCCCAG CTCGAACCGC AGCACCCCAG gtgagtagag
301 ggggagctgg aagaaggaag agagcggagc caggtctgtc actcgggcct
351 ctgcaagggtt tgtgatgtct tgaagtgccg agtgtcatta gatgtctgaa
401 ggcaagttag agccagcacc gcaagcaagt tgtgcgtgtg tgtcgggtgtg
451 tctgtgccgg tgtctcctca tcgtctggcc agtgagaatg aatgtctgtg
501 ggttcacctc tgtgtccacc cgacgacagg tgtgtgtaca tatgtatcct
551 gctctcagaa aatgggccta tgccgccggg cgcggtgact cacgcctgta
601 atcccaacac tgggaggctg aggcaggcag attacctgag gtcaggagtt
651 cgagaccagc caggccaaca tggggaaact ctgtctctac taaaaataaa
701 aattagcagg gcgtggtggc gggcgccctgt agtcccaact actcgggagg
751 ctgaggcagg agaatctctt gaacctggga ggcggagggtt gcagtcaagc
801 cgagatcaca ccactgcact ccagccaggg caacagagcg agatgcgtct
851 caaaaaaaaa aaaaaaaaaa aaaaggagag aaaacaaaaa gaaaagaaag
901 gaaaataggc ctatgccttc ctcagggtgtg tgctggggat ggtgggtgtt
951 acatcttcca agtctgggcc tgtgtctgtg ttggtgctcc ctgtcccaca
1001 tccagaaatc aagaagcgag ggctgggcag cagatatata gggtgagaag

Fig. 1

2/13

1051 ggaaggattt catgcattgt tacagtgatg cctggctgac ccttctcttt
EXON 2
1101 ccatcccaga TCCTAGCTGG GAGCCTGAAG GCTCCACTCT GGCTTCGTGC
1151 TTA CTTCAG GGCCTGCTCT TCTCTCTGGG ATGCGGGATC CAGAGACATT
1201 GTGGCAAAGT GCTCTTTCTG GGACTGTTGG CCTTTGGGGC CCTGGCATTA
1251 GGTCTCCGCA TGGCCATTAT TGAGACAAAC TTGGAACAGC TCTGGGTAGA
1301 AGTGGGCAGC CGGGTGAGCC AGGAGCTGCA TTACACCAAG GAGAAGCTGG
1351 GGGAGGAGGC TGCATACACC TCTCAGATGC TGATACAGAC CGCACGCCAG
1401 GAGGGAGAGA ACATCCTCAC ACCCGAAGCA CTTGGCCTCC ACCTCCAGGC
1451 AGCCCTCACT GCCAGTAAAG TCCAAGTATC ACTCTATGGG AAG.....
1501g
1551 tgagtctggc tgagcccctg agcagctggg ggcgaggcgt gctgtggggg
1601 ttctggagtg ggaatcccct tcttctgctg atctcctatg cccctggcta
EXON 4
1651 ttgcagTCCT GGGATTGAA CAAAATCTGC TACAAGTCAG GAGTTCCCCT
1701 TATTGAAAAT GGAATGATTG AGCGGgtaag tgtcctgaga gggagtagag
1751 gcagaacttt ttctgtagcg tgggaggact cagagaccga gcaagcccca
1801 cagcctgcaa tctgccccct taaaactaag gagggggatt gcagagggca
1851 tcctacaaag gttgtggggc aggactgacg tggcccgggg tatccctggc
EXON 5
1901 agATGATTGA GAAGCTGTTT CCGTGCGTGA TCCTCACCCC CCTCGACTGC
1951 TTCTGGGAGG GAGCCAAACT CCAAGGGGGC TCCGCCTACC TGCCgtgagt
2001 gccactcctg gggccctgct tcctctcccg ctggggactc tcccagcaga
2051 aaggaggggt ctggggaatg aggatgatca aaaccttacc aaggtcctaa
2101 ttacctccca ggccaggaac agagagcatg ggcttcccca aggctctctc
2151 cacatcctcc ttctctttcc ctctcaagga aggaagacct gacttattta
2201 cacaaaacta aacacaaaga tctgtaagat ctgagcaaag gagaaaaaga
2251 tccccacaaa gaggttttgc tgggggaaat tacctaggtg ttgctaagc
2301 cattgcccag gccagaaaga aaacctgcta caggcatgtg cctgctggtt
2351 gtatattaga accaagcaca cagcttggtg aggaactcag tggggccttt

Fig. 1 (cont.)

3/13

2401 ctgggccctt tctatgtatt aggtaaccct gccctgatat tcgtctcagc
2451 cccttgtagt cttctacagc tcaactgtagc accctgggtg gcccatgcag
2501 cctggcagtt ctgagaagct gaggcttgca caccctccat atggaaggac
2551 aaatcggcag ataagaggag ggtggggtag agcatggcgc cccagcagca
2601 gtttggagcc tgggttttcg tccctgaccc tcaccaacta taggcttttc
2651 cctcagCGGC CGCCCGGATA TCCAGTGGAC CAACCTGGAT CCAGAGCAGC
2701 TGCTGGAGGA GCTGGGTCCC TTTGCCTCCC TTGAGGGCTT CCGGGAGCTG
2751 CTAGACAAGG CACAGGTGGG CCAGGCCTAC GTGGGGCGGC CCTGTCTGCA
2801 CCCTGATGAC CTCCACTGCC CACCTAGTGC CCCCAACCAT CACAGCAGGC
2851 AGgtgggttc caaccaggtc tgccagggaa aggetgtttt ccttcctttt
2901 cccttcctca tactcctgtg ttctggggga gctgactgct ctgtgccctg
2951 accccccact tcctggccat tattaccctg ctcccacagt gccaggcccc
3001 caatgttcca ttccattca gttatcctac ggagccctca agtggtatat
3051 atgaatccct ttttcctttt ctaagcctag ataaggctgg acttcttttt
3101 tttttttttt ttgagtctca ctctgtcacc caggctggag tgcagtagtt
3151 cgatcttggc tcaactgcaac ctcggtcaa gcaattctcc tgccttagcc
3201 tcctgagtag ctgggattac aggtgcccac caccatgccc ggctaatttt
3251 tattagcctc ccaaagtgtc gggattacag gcgtgagcca ctgcgcctgg
3301 ccaaggctgg actttttatc aaaatagact aatacagggga aactaagaac
3351 acagcaggta agcatgaata tcatacctgg tttcccaggt ttctttgtgg
3401 ccctgcaa at gtggtacttt tttcagaatc cgccagttac accagctcct
3451 cccagaagcc tacttccagg cctctgcttc cccttggggc ttctgtctg
3501 cgggatacta gctgttact cctgcagagc agtcaagagg ctccagaatag
3551 ttacctacac tccagcccta ctgagcttca tggcagcgtg gttcctggag
3601 gtggaagccc agggacactc agttatccac ggccagggcc ttgagcatta
3651 acccctcctg ttccccctca gGGCTCCCAA TGTGGCTCAC GAGCTGAGTG
3701 GGGGCTGCCA TGGCTTCTCC CACAAATTCA TGCACTGGCA GGAGGAATTG

Fig. 1 (cont.)

3751 CTGCTGGGAG GCATGGCCAG AGACCCCAAG GGAGAGCTGC TGAGgtaggg
3801 tctcctctgg gagttggtga ggggactctg ttcattgagaa cccatactgt
3851 aatgccaggc agctctggca aaaggccctt cacatccctc accagggtgtt
EXON 8
3901 tggggccagct ctgacccttg gttctccac acccccacca gGGCAGAGGC
3951 CCTGCAGAGC ACCTTCTTGC TGATGAGTCC CCGCCAGCTG TACGAGCATT
4001 TCCGGGGTGA CTATCAGACA CATGACATTG GCTGGAGTGA GGAGCAGGCC
4051 AGCACAGTGC TACAAGCCTG GCAGCGGCGC TTTGTGCAGg tcggtatgga
4101 caaggacaag gggggtgccc tgaggccatt ccctcctcct gccccctcct
EXON 9
4151 atccaccctg tttctccagc TGGCCCAGGA GGCCCTGCCT GAGAACGCTT
4201 CCCAGCAGAT CCATGCCTTC TCCTCCACCA CCCTGGATGA CATCCTGCAT
4251 GCGTTCTCTG AAGTCAGTGC TGCCCGTGTG GTGGGAGGCT ATCTGCTCAT
4301 Ggtgggtctt gcacctggca ccttgcccc accccacctc caaccagtgc
EXON 10
4351 ccaccctggg agccctgag actgcccctt cccccacag CTGGCCTATG
4401 CCTGTGTGAC CATGCTGCGG TGGGACTGCG CCCAGTCCCA GGGTTCCGTG
4451 GGCCTTGCCG GGGTACTGCT GGTGGCCCTG GCGGTGGCCT CAGGCCTTGG
4501 GCTCTGTGCC CTGCTCGGCA TCACCTTCAA TGCTGCCACT ACCCAGgtac
4551 gccaggactg cagggcagac tcagtgccag tcaccagggt tcacgggtcc
EXON 11
4601 tcagtgtccc gctcctctgc ccctccagGT GCTGCCCTTC TTGGCTCTGG
4651 GAATCGGCGT GGATGACGTA TTCCTGCTGG CGCATGCCTT CACAGAGGCT
4701 CTGCCTGGCA CCCCTCTCCA Ggtggggcct tgtccccag ggctcatctg
4751 aggcagctca gcttactggt taagagcctc ttggttcaag tgacccttgg
4801 gctgctaata aacctcgggt cctcttgctc ccatctgtaa acaggggaaa
4851 taatagtgt gtgtcctaag gggtattggt tggatcagtg aggtaactca
4901 agttgaatgc ttagaacagc ccatcatag tacatggtac ccaataaatg
4951 ctagccactg tggtatgact gccccacctc tgcaccccaa gttcctgagc
5001 ctccccttca ctccactttg acacggcccc tcccttggtga cctgagggca
EXON 12
5051 ggtccccact ctgtcctggc agGAGCGCAT GGGCGAGTGT CTGCAGCGCA

Fig. 1 (cont.)

5/13

EXON 13

EX

XON 14

EXON 15

Fig. 1 (cont.)

6/13

6451 gagagggctg gaggggtcca ctagtacagg ggctgcaggc ctcctgggcc
EXON 16
6501 caggccttca gccctctctg cctctgcagC TGACCACAAG GAAGCTGGTG
6551 GACAGAGAGG GACTGATTCC ACCCGAGCTC TTCTACATGG GGCTGACCGT
6601 GTGGGTGAGC AGTGACCCCC TGGGTCTGGC AGCCTCACAG GCCAACTTCT
6651 ACCCCCCACC TCCTGAATGG CTGCACGACA AATACGACAC CACGGGGGAG
6701 AACTTTCGCA gtgagtcttg gggggagctc ggcaagagcc tcagcctcgc
6751 ccacacaagc cctgagcctg aggccttgc cactctgccc cgtgctcacc
EXON 17
6801 gccctgtccc tctccctctt ctcccttccc ctccccctca cagTCCCGCC
6851 AGCTCAGCCC TTGGAGTTTG CCCAGTTCCC TTTCTGCTG CGTGGCCTCC
6901 AGAAGACTGC AGACTTTGTG GAGGCCATCG AGGGGGCCCG GGCAGCATGC
6951 GCAGAGGCCG GCCAGGCTGG GGTGCACGCC TACCCCAGCG GCTCCCCCTT
7001 CCTCTTCTGG GAACAGTATC TGGGCCTGCG GCGCTGCTTC CTGCTGGCCG
7051 TCTGCATCCT GCTGGTGTGC ACTTTCCTCG TCTGTGCTCT GCTGCTCCTC
7101 AACCCCTGGA CGGCTGGCCT CATAgtgagt gcttgcagga gtggggacag
7151 agacacccca cccttccctg cccagcctgt catccctcct gccaggagcc
EXON 18
7201 ctctgtgagc cctgtctccc tcagGTGCTG GTCCTGGCGA TGATGACAGT
7251 GGAACTCTTT GGTATCATGG GTTTCCTGGG CATCAAGCTG AGTGCCATCC
7301 CCGTGGTGAT CTTGTGGCC TCTGTAGGCA TTGGCGTTGA GTTCACAGTC
7351 CACGTGGCTC TGGGCTTCCT GACCACCCAG GGCAGCCGA ACCTGCGGGC
7401 CGCCCATGCC CTTGAGCACA CATTTGCCCC CGTGACCGAT GGGGCCATCT
7451 CCACATTGCT GGGTCTGCTC ATGCTTGCTG GTTCCCACTT TGACTTCATT
7501 GTAAG.....
7551 gtagggaggg ctcggggcag ggaggcaggg ctcaggacag
EXON 20
7601 gcctgggctg actccccca caccctaccc ctagGTAATT CTTTGCGGCG
7651 CTGACAGTGC TCACGCTCCT GGGCCTCCTC CATGGACTCG TGCTGCTGCC
7701 TGTGCTGCTG TCCATCCTGG GCCCGCCGCC AGAGgtgacc acaccctcgg
7751 caccatccct ctactcccag cccaagggac ggggtaggga gaggcaaggg

Fig. 1 (cont.)

7/13

7801 aagggacaga gccctgtggc ccacagacag gtacctcccc aacaggtgcc
7851 accagctgaa ggtggcagcc tcctcctttc cccagacacc atgttcctgc
7901 ccctcagccc tcctggcttc ttcattgggac ccaccttaga cttttaggat
7951 ccagaacaag gtgcaggggt tgccccaggc ctcaacatcc tgtcgcctgc
8001 cagctctcat atcctgctgg agaccaacaa gggccccagc ttcccaacag
8051 tcatggtaat cccagcagag atgctaaagg ggacgggagc cccagggggc
8101 cgtgggctta ctggggctgg tgtctcccca cagGTGATAC AGATGTACAA
8151 GGAAAGCCCA GAGATCCTGA GTCCACCAGC TCCACAGGGA GCGGGGCTTA
8201 Ggtggggggc atcctcctcc ctgccccaga gctttgccag agtgactacc
8251 tccatgaccg tggccatcca cccaccccc ctgcctgggtg cctacatcca
8301 tccagcccct gatgagcccc cttgggtcccc tgctgtcact agctctggca
8351 acctcagttc cagggggacca ggtccagcca ctgggtgaaa gagcagctga
8401 agcacagaga ccatgtgtgg ggcgtgtggg gtcactggga agcactgggt
8451 ctggtgtag acgcaggatg gaccctgga gggctctgct gctgctgcat
8501 cccctctccc gaccagctg tcatgggcct ccctgatatc catacagaac
8551 agccaccgat ttgcacatcc aggctgtgt gagcctgtat ctgtgtcact
8601 tgagagtga agctggcact tggggctgca gtgcagccct gtcccccttc
8651 ccaccccaca ccaactgctg cccagctgac caagcctgag ggaccctcca
8701 gcacccttcc gtctgggtgac tcctgggcag gctctccata tccctgcccc
8751 cctcctacca catccattat ttatatgaaa atgtctattt ttgtagtata
8801 catacatgtt agctatgatg aaagttttat tttttaaaga atgaaatata
8851 ttctatgtga agctatgatg aaagttttat tttttaaaga atgaaatata
8901 ttctatgtga actaatctcg aaagttttat tttttaaaga atgaaatata
8951 ttctatgtgt gcaagtgaac attagcttca gttgcttttt tttggacaga
9001 gtggggagtt tgcaagtga cattaactat tggaaggagc ttctctgggtg
9051 ccaggacctg aggtattagc ttctctagtt ctgggtggaa aagaccccag
9101 attctggatt tttgtcatac acttggtaac atcatctgga ttaagtgcct

Fig. 1 (cont.)

8/13

9151 actatacaaaa acgataacaa attttggttg tgtgaaatcc tactgggttc
9201 aatctggaga ccgagagcag aaaaaaaga accccactgt gtggctttca
9251 gagccaccat attccagcct gcccgctctt ccagactcac ctccacctac
9301 ctgcttcacc cgcacgggaa acggcaaggc agaggggcaa agccatgcag
9351 caggtggaag gcgaggtgga ggcagatcag gaaagcagcc agttgaagca
9401 gagagaggtc aacaggttct ggggagcttc tcaggaggtt tgtggacca
9451 gggaaaggag ccaggttcca gagcaacctc caaggcaaag gcctctgtaa
9501 gttggttgct ctgacagccg agaggtgtct ttggccagtc agccagtgga
9551 tcagttgcgg gaactgctca gaaactgagg tgctagcagt tagtgaggac
9601 acagcgtaag ttgtttgttc tgtgaaagtt gaacagctcc actaagcaga
9651 ggccttgaag agtggccaca gccctggaat agagcacaga gcctcaccta
9701 gaggcgtggg gaggtttgca actgcccctt cccagccata gcttaggacc
9751 catagtctag ttcacataga ccctgggctc caaccacca ctcaccagga
9801 atgatccac cccaggaaca atgcgttctc acatcccacc ccacctggac
9851 aaaggccagg aaatcatgtt ctgaccaaaa gatacaacaa caaaaacaac
9901 aacaacaaaa aacgcctatt gcaattgaat ccacgctaaa atgcctaaaa
9951 agctcaagag aagcgggtag ttggcagaga acctagagta gggggtgcaa
10001 ccagcaggcc caaggagggg aggctgcatt tgggtccagc agtgtttggg
10051 tcaccaagaa gggccttcta ggtggagcag agagagctca ccaggccaga
10101 atagtcaaaa gggggtcagc cctcagtgcc acttaccagc ggagtaaccc
10151 tgggcaagtt agccagcctc actaagcctc cccatcttca tctttccagG
XON 22
10201 CCCGAGGAGA TC**TAG**CTCT GCCTCCCACC CCAGCACCCC CTCATCAGAC
10251 ACAAGGAGCG CCACTGTCTG GACAGGCTGA ATTGGTCTTC GGGTCCCTAA
10301 TTTCTCATAC GCCATTCCCT CTGCCTAGAA CACTTTCTCA CCTCCCCTTG
10351 ATGTGACCCC ATATCACCTT TCGAGGTGAA TTGGATCGGA TGCCATCTCC
10401 TCCAGGAGGG GTGGGGTCGT GCCTCCTGTG AGGTCCCAGT GCCCCTGAGT
10451 GTCTGTGCCC GTCTGTTTCC CCGTCCCTCT CTCTAAGCCC GGAGGCTTAC

Fig. 1 (cont.)

9/13

10501 TCGGGGTAAG GACGGCGGGA CAGGACCTTA ACCCCTGGGA CGAACACCAG
10551 CTCCGCAAAG GACTCCGCAC CCGGCGCCGC CCACGGGGTG CGGGTCCCAG
10601 GAGGACCAGC AGAGAGGAGC ATAGGAGAGC AAAGGAGATC AGTGACCCAT
10651 GGCTTCCCCG GTGGCGCGGA ACAGCCCGGA GCCGCCTGTG ATTTGCATAC
10701 CCATGGTGCA CCACGAAAAG ATACCCTCAA GATGCTTGCA CTCCTCTGT
10751 GCGCGCATTT CTGCACTGTT TTAGAGCATG ATGCCTCTTA CACGCATCTG
10801 TGTGCATAAA CTACATATAG GGAGTGCGTA CCACGCAGGC ATCCAACAAC
10851 CATAAGTGTG TTAAGTGTTA GTTCTCCCTG CGAGGTTCGA AGCGGAAGTC
10901 ACGAATATAC TCGGGTTTCT CTTCAAAGCG CATAAATCTT TCGCCTTTTA
10951 CTAAAGATTT CCGTGGAGAG AAAGTTGTGA GTTTTTATTC AATTTTTTGA
11001 GGCCTCTTAT TTCCTGAGGC TACATTTTAA AGTATTAAAA GTTAGGCAAC
11051 TACAAAAAAA AAAAAAAA

Fig. 1 (cont.)

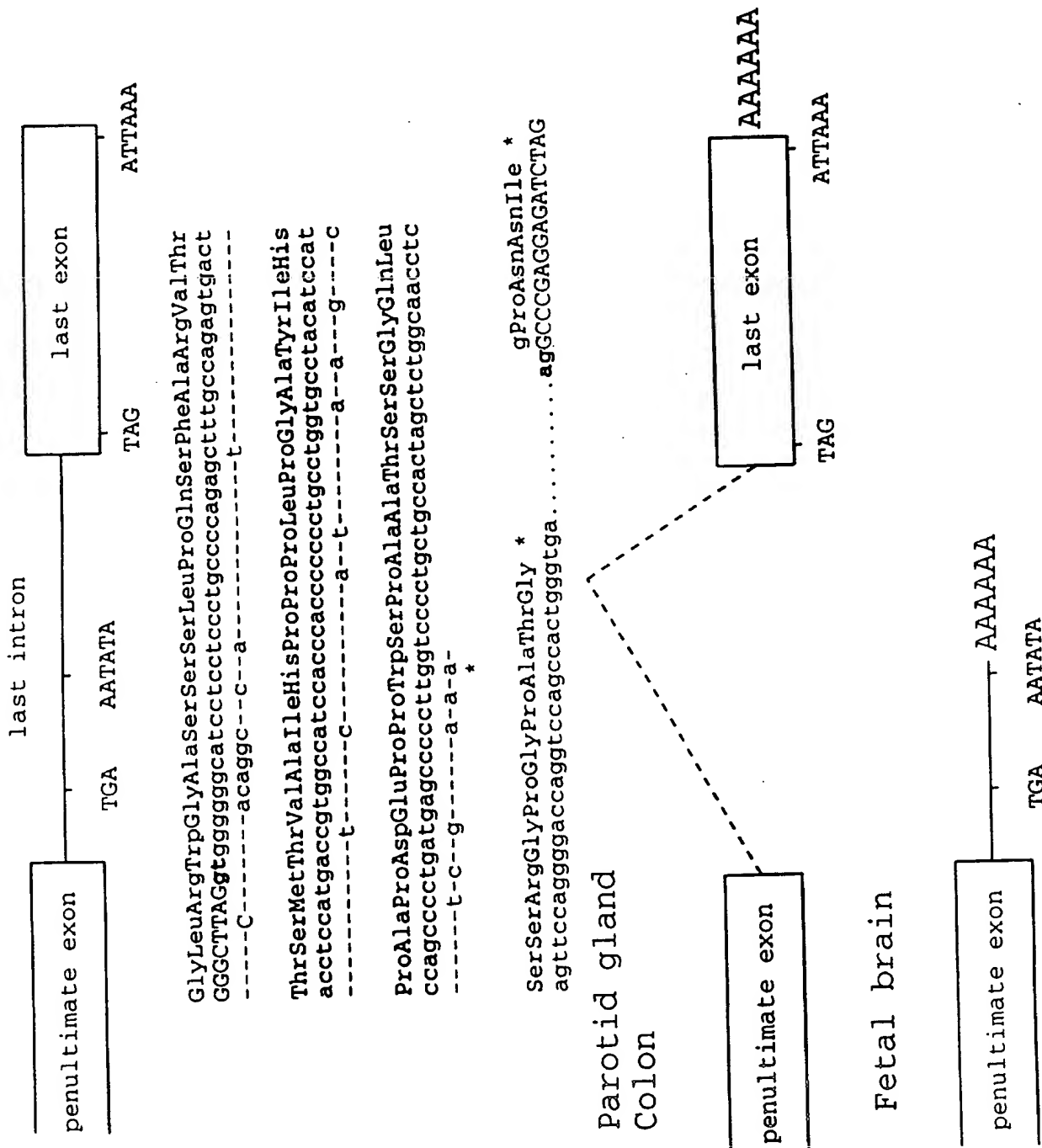
10/13

1MTRSPPLREL P..... 11
 1 MASAGNAAEPQDRGGGGSGCIGAPGRPAGGGRRRRRTGGLRRAAAPDRDYL 50
 12 ..PSYTPPARTAAPQI...LAGSLKAPLWLRAYFQGLLFSLGCGIQRHCG 56
 51 HRPSYCDAA.FALEQISKGKATGRKAPLWLRKFQRLFLKLGCIYQKNCG 99
 57 KVLFLGLLAFGALALGLRMAIIETNLEQLWVEVGSRSVSELHYTKEKLGE 106
 100 KFLVVGLLIFGAFVGLKAAANLETNVEELWVEVGGRVSRELNYTRQKIGE 149
 107 EAAYTSQMLIQATARQEGENILTPEALGLHLQAALTASKVQVSLYGKSWDL 156
 150 EAMFNPQLMIQTPKEEGANVLTTEALLQHLDSALQASRVHVVMYNRQWKL 199
 157 NKICYKSGVPLIENGMIERMIEKLFPCVILTPLDCFWEAGAKLQGG SAYLP 206
 200 EHLCKYKSGELITETGYMDQIIEYLYPCLIIITPLDCFWEAGAKLQSGTAYLL 249
 207 GRPDIQWTNLDPEQLLEELGPFA.SLEGFRELLDKAQVGQAYVGRPCLHP 255
 250 GKPPLRWTFNDFLEFLEELKKINYQVDSWEEMLNKAEVGHGYMDR PCLNP 299
 256 DDLHCPPSAPNHHSRQAPNVAAHELSGGCHGFSHKFMHWQEELLLGGMARD 305
 300 ADPDCPATAPNKNSTKPLDMALVLNGGCHGLSRKYMHWQEELIVGGTVKN 349
 306 PQGELLRAEALQSTFLLMSPROLYEHFRGDYQTHDIGWSEEQASTVLQAW 355
 350 STGKLVSAAHALQTMFQLMTPKQMYEHFKGYEYVSHINWNEDKAAAIL EAW 399
 356 QRRFVQLAQEALPENASQQIHAFSSTLDDILHAFSEVSAARVVG GYLLM 405
 400 QRTYVEVVHQSV AQNSTQKVLSTTTTLLDILKSFSVSVIRVASGYLLM 449
 406 LAYACVTMLRWDCAQSQSGVLAGVLLVALAVASGLGLCALLGITFNAAT 455
 450 LAYACL TMLRWDCSKSQGAVGLAGVLLVALSVAAGLGLCSLIGISFNAAT 499
 456 TQVLPFLALGIGVDDVFLLAHAFTEALPG..TPLQERMGECLQRTGTSVV 503
 500 TQVLPFLALGVGVDDVFLLAHAFSETGQNKRI PFEDRTGECLKRTGASVA 549
 504 LTSINNMAAFLMAALVPIPALRAFSLQAAIIVVGCTFVAVMLVFPAILSLD 553
 550 LTSISNVTAFFMAALIPIPALRAFSLQAAVVVVFNFAMVLLIFPAIL SMD 599
 554 LRRRHCRQLDVLCCFSSPCSAQVIQILPQELGDGT.....VPVG 592
 600 LYRREDRRLDIFCCFTSPCVSRVIOVEPQAYTDTHDNTRYSPPPPYSSH S 649
 593 IAH.....LTATVQAFTHCEASSQHVV TILPPQ AHL....VPPPSDPLGS 633
 650 FAHETQITMQSTVQLRTEYDPHTHVYYTTAEPRSEISVQPVTVTQDTL SC 699

Fig. 2A

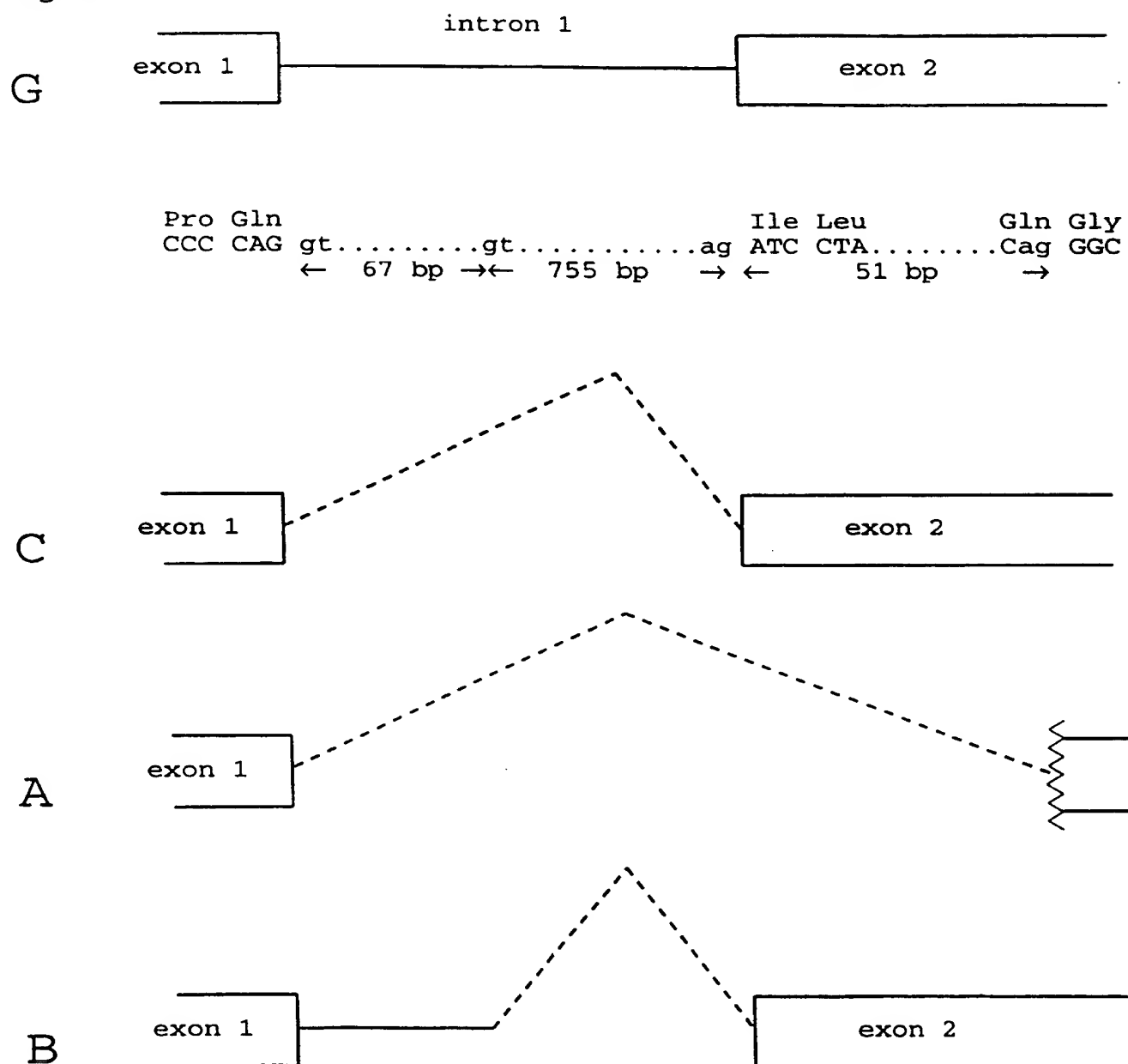
11/13

Fig. 2B Genomic



12/13

Fig. 2C



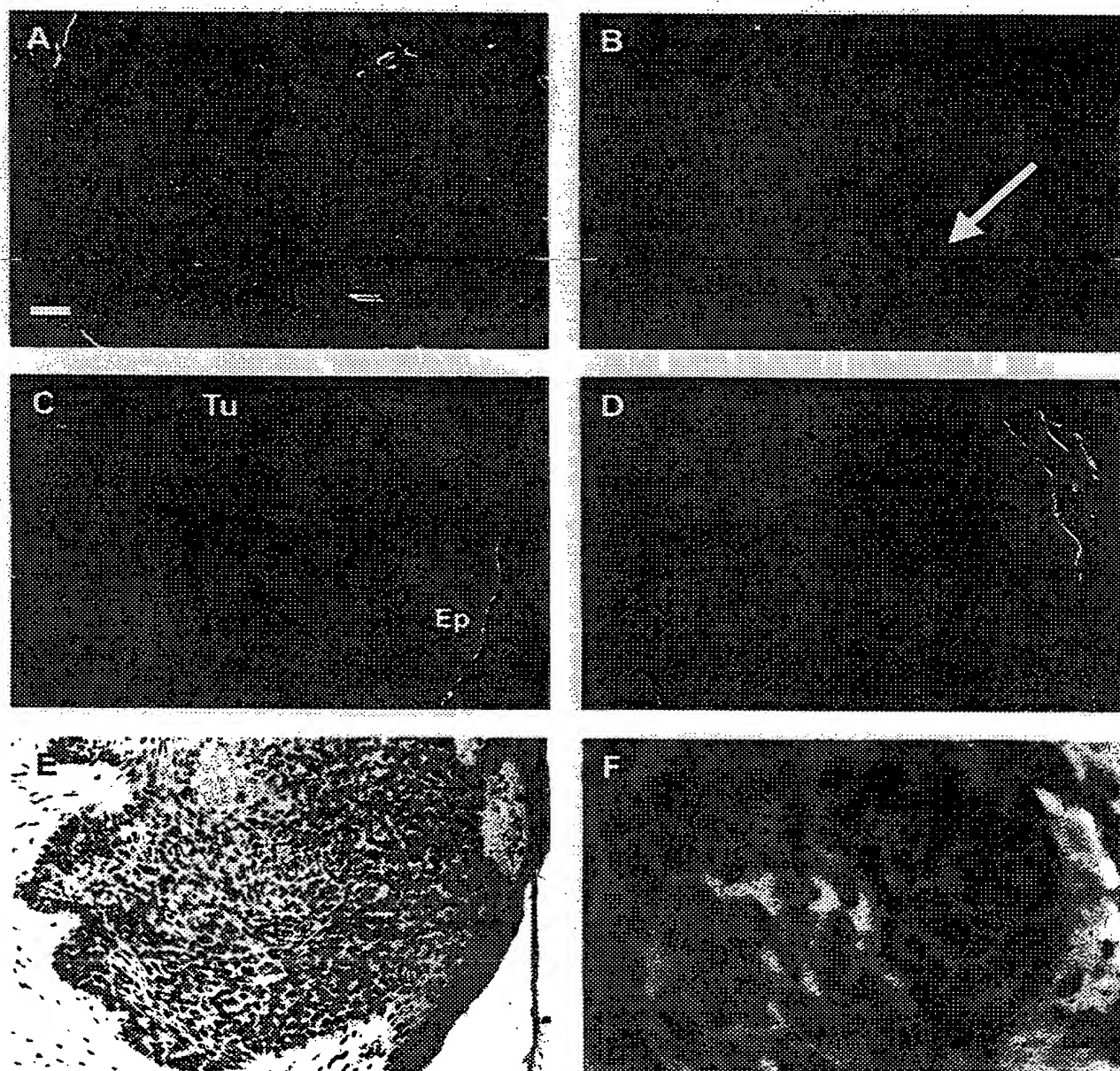


Fig. 3

SEQUENCE LISTING

<110> PHARMACIA & UPJOHN S.P.A.
KAROLINSKA INNOVATIONS AB

<120> A NOVEL COMPONENT IN THE HEDGEHOG SIGNALLING PATHWAY

<130> 56822

<140>

<141>

<160> 5

<170> PatentIn Ver. 2.1

<210> 1

<211> 1146

<212> PRT

<213> HUMAN

<400> 1

```

Met Thr Arg Ser Pro Pro Leu Arg Glu Leu Pro Pro Ser Tyr Thr Pro
 1              5              10              15

Pro Ala Arg Thr Ala Ala Pro Gln Ile Leu Ala Gly Ser Leu Lys Ala
      20              25              30

Pro Leu Trp Leu Arg Ala Tyr Phe Gln Gly Leu Leu Phe Ser Leu Gly
 35              40              45

Cys Gly Ile Gln Arg His Cys Gly Lys Val Leu Phe Leu Gly Leu Leu
 50              55              60

Ala Phe Gly Ala Leu Ala Leu Gly Leu Arg Met Ala Ile Ile Glu Thr
 65              70              75              80

Asn Leu Glu Gln Leu Trp Val Glu Val Gly Ser Arg Val Ser Gln Glu
      85              90              95

Leu His Tyr Thr Lys Glu Lys Leu Gly Glu Glu Ala Ala Tyr Thr Ser
      100              105              110

Gln Met Leu Ile Gln Thr Ala Arg Gln Glu Gly Glu Asn Ile Leu Thr
      115              120              125

Pro Glu Ala Leu Gly Leu His Leu Gln Ala Ala Leu Thr Ala Ser Lys
      130              135              140

Val Gln Val Ser Leu Tyr Gly Lys Ser Trp Asp Leu Asn Lys Ile Cys
      145              150              155              160

Tyr Lys Ser Gly Val Pro Leu Ile Glu Asn Gly Met Ile Glu Arg Met
      165              170              175

Ile Glu Lys Leu Phe Pro Cys Val Ile Leu Thr Pro Leu Asp Cys Phe
      180              185              190

Trp Glu Gly Ala Lys Leu Gln Gly Gly Ser Ala Tyr Leu Pro Gly Arg
      195              200              205

```

Pro Asp Ile Gln Trp Thr Asn Leu Asp Pro Glu Gln Leu Leu Glu Glu
 210 215 220
 Leu Gly Pro Phe Ala Ser Leu Glu Gly Phe Arg Glu Leu Leu Asp Lys
 225 230 235 240
 Ala Gln Val Gly Gln Ala Tyr Val Gly Arg Pro Cys Leu His Pro Asp
 245 250 255
 Asp Leu His Cys Pro Pro Ser Ala Pro Asn His His Ser Arg Gln Ala
 260 265 270
 Pro Asn Val Ala His Glu Leu Ser Gly Gly Cys His Gly Phe Ser His
 275 280 285
 Lys Phe Met His Trp Gln Glu Glu Leu Leu Leu Gly Gly Met Ala Arg
 290 295 300
 Asp Pro Gln Gly Glu Leu Leu Arg Ala Glu Ala Leu Gln Ser Thr Phe
 305 310 315 320
 Leu Leu Met Ser Pro Arg Gln Leu Tyr Glu His Phe Arg Gly Asp Tyr
 325 330 335
 Gln Thr His Asp Ile Gly Trp Ser Glu Glu Gln Ala Ser Thr Val Leu
 340 345 350
 Gln Ala Trp Gln Arg Arg Phe Val Gln Leu Ala Gln Glu Ala Leu Pro
 355 360 365
 Glu Asn Ala Ser Gln Gln Ile His Ala Phe Ser Ser Thr Thr Leu Asp
 370 375 380
 Asp Ile Leu His Ala Phe Ser Glu Val Ser Ala Ala Arg Val Val Gly
 385 390 395 400
 Gly Tyr Leu Leu Met Leu Ala Tyr Ala Cys Val Thr Met Leu Arg Trp
 405 410 415
 Asp Cys Ala Gln Ser Gln Gly Ser Val Gly Leu Ala Gly Val Leu Leu
 420 425 430
 Val Ala Leu Ala Val Ala Ser Gly Leu Gly Leu Cys Ala Leu Leu Gly
 435 440 445
 Ile Thr Phe Asn Ala Ala Thr Thr Gln Val Leu Pro Phe Leu Ala Leu
 450 455 460
 Gly Ile Gly Val Asp Asp Val Phe Leu Leu Ala His Ala Phe Thr Glu
 465 470 475 480
 Ala Leu Pro Gly Thr Pro Leu Gln Glu Arg Met Gly Glu Cys Leu Gln
 485 490 495
 Arg Thr Gly Thr Ser Val Val Leu Thr Ser Ile Asn Asn Met Ala Ala
 500 505 510
 Phe Leu Met Ala Ala Leu Val Pro Ile Pro Ala Leu Arg Ala Phe Ser
 515 520 525
 Leu Gln Ala Ala Ile Val Val Gly Cys Thr Phe Val Ala Val Met Leu

530					535					540					
Val 545	Phe	Pro	Ala	Ile	Leu 550	Ser	Leu	Asp	Leu	Arg 555	Arg	Arg	His	Cys	Gln 560
Arg	Leu	Asp	Val	Leu 565	Cys	Cys	Phe	Ser	Ser 570	Pro	Cys	Ser	Ala	Gln 575	Val
Ile	Gln	Ile	Leu 580	Pro	Gln	Glu	Leu	Gly 585	Asp	Gly	Thr	Val	Pro 590	Val	Gly
Ile	Ala	His 595	Leu	Thr	Ala	Thr	Val 600	Gln	Ala	Phe	Thr	His 605	Cys	Glu	Ala
Ser	Ser 610	Gln	His	Val	Val	Thr 615	Ile	Leu	Pro	Pro	Gln 620	Ala	His	Leu	Val
Pro 625	Pro	Pro	Ser	Asp	Pro 630	Leu	Gly	Ser	Glu	Leu 635	Phe	Ser	Pro	Gly	Gly 640
Ser	Thr	Arg	Asp	Leu 645	Leu	Gly	Gln	Glu	Glu 650	Glu	Thr	Arg	Gln	Lys 655	Ala
Ala	Cys	Lys	Ser 660	Leu	Pro	Cys	Ala	Arg 665	Trp	Asn	Leu	Ala	His 670	Phe	Ala
Arg	Tyr	Gln 675	Phe	Ala	Pro	Leu	Leu 680	Leu	Gln	Ser	His	Ala 685	Lys	Ala	Ile
Val 690	Leu	Val	Leu	Phe	Gly	Ala 695	Leu	Leu	Gly	Leu	Ser 700	Leu	Tyr	Gly	Ala
Thr 705	Leu	Val	Gln	Asp	Gly 710	Leu	Ala	Leu	Thr	Asp 715	Val	Val	Pro	Arg	Gly 720
Thr	Lys	Glu	His 725	Ala	Phe	Leu	Ser	Ala	Gln 730	Leu	Arg	Tyr	Phe	Ser 735	Leu
Tyr	Glu	Val	Ala 740	Leu	Val	Thr	Gln	Gly 745	Gly	Phe	Asp	Tyr	Ala 750	His	Ser
Gln	Arg	Ala 755	Leu	Phe	Asp	Leu	His 760	Gln	Arg	Phe	Ser	Ser 765	Leu	Lys	Ala
Val 770	Leu	Pro	Pro	Pro	Ala	Thr 775	Gln	Ala	Pro	Arg	Thr 780	Trp	Leu	His	Tyr
Tyr 785	Arg	Asn	Trp	Leu	Gln 790	Gly	Ile	Gln	Ala	Ala 795	Phe	Asp	Gln	Asp	Trp 800
Ala	Ser	Gly	Arg	Ile 805	Thr	Arg	His	Ser	Tyr 810	Arg	Asn	Gly	Ser	Glu 815	Asp
Gly	Ala	Leu	Ala 820	Tyr	Lys	Leu	Leu	Ile 825	Gln	Thr	Gly	Asp	Ala 830	Gln	Glu
Leu	Leu	Asp 835	Phe	Ser	Gln	Leu	Thr 840	Thr	Arg	Lys	Leu	Val 845	Asp	Arg	Glu
Gly	Leu 850	Ile	Pro	Pro	Glu	Leu 855	Phe	Tyr	Met	Gly	Leu 860	Thr	Val	Trp	Val

Ser Ser Asp Pro Leu Gly Leu Ala Ala Ser Gln Ala Asn Phe Tyr Pro
 865 870 875 880
 Pro Pro Pro Glu Trp Leu His Asp Lys Tyr Asp Thr Thr Gly Glu Asn
 885 890 895
 Phe Arg Ile Pro Pro Ala Gln Pro Leu Glu Phe Ala Gln Phe Pro Phe
 900 905 910
 Leu Leu Arg Gly Leu Gln Lys Thr Ala Asp Phe Val Glu Ala Ile Glu
 915 920 925
 Gly Ala Arg Ala Ala Cys Ala Glu Ala Gly Gln Ala Gly Val His Ala
 930 935 940
 Tyr Pro Ser Gly Ser Pro Phe Leu Phe Trp Glu Gln Tyr Leu Gly Leu
 945 950 955 960
 Arg Arg Cys Phe Leu Leu Ala Val Cys Ile Leu Leu Val Cys Thr Phe
 965 970 975
 Leu Val Cys Ala Leu Leu Leu Leu Asn Pro Trp Thr Ala Gly Leu Ile
 980 985 990
 Val Leu Val Leu Ala Met Met Thr Val Glu Leu Phe Gly Ile Met Gly
 995 1000 1005
 Phe Leu Gly Ile Lys Leu Ser Ala Ile Pro Val Val Ile Leu Val Ala
 1010 1015 1020
 Ser Val Gly Ile Gly Val Glu Phe Thr Val His Val Ala Leu Gly Phe
 1025 1030 1035 1040
 Leu Thr Thr Gln Gly Ser Arg Asn Leu Arg Ala Ala His Ala Leu Glu
 1045 1050 1055
 His Thr Phe Ala Pro Val Thr Asp Gly Ala Ile Ser Thr Leu Leu Gly
 1060 1065 1070
 Leu Leu Met Leu Ala Gly Ser His Phe Asp Phe Ile Val Arg Tyr Phe
 1075 1080 1085
 Phe Ala Ala Leu Thr Val Leu Thr Leu Leu Gly Leu Leu His Gly Leu
 1090 1095 1100
 Val Leu Leu Pro Val Leu Leu Ser Ile Leu Gly Pro Pro Pro Glu Val
 1105 1110 1115 1120
 Ile Gln Met Tyr Lys Glu Ser Pro Glu Ile Leu Ser Pro Pro Ala Pro
 1125 1130 1135
 Gln Gly Gly Gly Leu Arg Pro Glu Glu Ile
 1140 1145

<210> 2
 <211> 3453
 <212> DNA
 <213> HUMAN

<400> 2

atcccccgcca	gcatgactcg	atcgccgccc	ctcagagagc	tgcccccgag	ttacacaccc	60
ccagctcgaa	ccgcagcacc	ccagatccta	gctgggagcc	tgaaggctcc	actctggctt	120
cgtgcttact	tccagggcct	gctcttctct	ctgggatgcg	ggatccagag	acattgtggc	180
aaagtgtctt	ttctgggact	gttggccttt	ggggccctgg	cattaggtct	ccgcatggcc	240
attattgaga	caaacttgga	acagctctgg	gtagaagtgg	gcagccgggt	gagccaggag	300
ctgcattaca	ccaaggagaa	gctgggggag	gaggctgcat	acacctctca	gatgctgata	360
cagaccgcac	gccaggaggg	agagaacatc	ctcacacccg	aagcacttgg	cctccacctc	420
caggcagccc	tactgcccag	taaagtccaa	gtatcactct	atgggaagtc	ctgggatttg	480
aacaaaatct	gctacaagtc	aggagttccc	cttattgaaa	atggaatgat	tgagcggatg	540
attgagaagc	tggttccgtg	cgtgatccct	acccccctcg	actgcttctg	ggagggagcc	600
aaactccaag	ggggctccgc	ctacctgccc	ggccgcccgg	atatccagtg	gaccaacctg	660
gatccagagc	agctgctgga	ggagctgggt	ccctttgcct	cccttgaggg	cttccgggag	720
ctgctagata	aggcacaggt	gggccaggcc	tacgtggggc	ggccctgtct	gcacctgat	780
gacctccact	gccccactag	tgcccccaac	catcacagca	ggcaggctcc	caatgtggct	840
cacgagctga	gtgggggctg	ccatggcttc	tcccacaaat	tcatgcaact	gcaggaggaa	900
ttgctgctgg	gaggcatggc	cagagacccc	caaggagagc	tgctgagggc	agaggccctg	960
cagagcacct	tcttgctgat	gagtccccgc	cagctgtacg	agcatttccg	gggtgactat	1020
cagacacatg	acattggctg	gagtggagg	caggccagca	cagtgtctaca	agcctggcag	1080
cggcgctttg	tgacgtggc	ccaggaggcc	ctgcctgaga	acgcttccca	gcagatccat	1140
gccttctcct	ccaccacct	ggatgacatc	ctgcatgcgt	tctctgaagt	cagtgtctgc	1200
cgtgtggtgg	gaggctatct	gctcatgctg	gcctatgcct	gtgtgacct	gctgcgggtg	1260
gactgcgccc	agtcccagg	ttccgtgggc	cttgccgggg	tactgctggt	ggccctggcg	1320
gtggcctcag	gccttgggct	ctgtgccctg	ctcggcatca	ccttcaatgc	tgccactacc	1380
caggtgctgc	ccttcttggc	tctgggaatc	ggcgtggatg	acgtattcct	gctggcgcat	1440
gccttcacag	aggctctgcc	tggcacccct	ctccaggagc	gcattggcga	gtgtctgagc	1500
cgcacgggca	ccagtgtcgt	actcacatcc	atcaacaaca	tggccgcctt	cctcatggct	1560
gcctcgttcc	ccatccctgc	gctgcgagcc	ttctccctac	aggcgcccat	agtgggtggc	1620
tgacaccttg	tagccgtgat	gcttgtcttc	ccagccatcc	tcagcctgga	cctacggcgg	1680
cgccactgcc	agcgccttga	tgtgctctgc	tgcttctcca	gtccctgctc	tgctcagggt	1740
attcagatcc	tgcccagga	gctgggggac	gggacagtac	cagtgggcat	tgcccacctc	1800
actgccacag	ttcaagcctt	tacccactgt	gaagccagca	gccagcatgt	ggtcaccatc	1860
ctgcctcccc	aagcccacct	ggtgccccca	ccttctgacc	caactgggctc	tgagctcttc	1920
agccctggag	ggtccacacg	ggaccttcta	ggccaggagg	aggagacaag	gcagaaggca	1980
gcctgcaagt	ccctgccttg	tgcccgtgg	aatcttgccc	atttcgcccg	ctatcagttt	2040
gccccgttgc	tgctccagtc	acatgctaag	gccatcgtgc	tggtgtctct	tggtgtctct	2100
ctgggcctga	gcctctacgg	agccaccttg	gtgcaagacg	gcctggccct	gacggatgtg	2160
gtgcctcggg	gcaccaagga	gcatgccttc	ctgagcgccc	agctcaggta	cttctccctg	2220
tacgaggtgg	ccctggtgac	ccagggtggc	tttgactacg	ccactccca	acgcgccttc	2280
tttgatctgc	accagcgctt	cagttccctc	aaggcggtgc	tgccccccacc	ggccaccagc	2340
gcaccccgca	cctggctgca	ctattaccgc	aactggctac	agggaatcca	ggctgccttt	2400
gaccaggact	gggttctctg	gcgcataccc	cgcactcgt	accgcaatgg	ctctgaggat	2460
ggggcccttg	cctacaagct	gctcatccag	actggagacg	cccaggagct	tctggatttc	2520
agccagctga	ccacaaggaa	gctggtggac	agagaggggc	tgattccacc	cgagctcttc	2580
tacatggggc	tgaccgtgtg	ggtgagcagt	gaccccttgg	gtctggcagc	ctcacaggcc	2640
aacttctacc	ccccacctcc	tgaatggctg	cacgacaaat	acgacaccac	gggggagaa	2700
tttcgcatcc	cgccagctca	gcccttggag	tttgccagct	tcccttctct	gctgcgtggc	2760
ctccagaaga	ctgcagactt	tgtggaggcc	atcgaggggg	cccgggcagc	atgcccagag	2820
gcccggcagg	ctgggggtgca	cgcctacccc	agcggctccc	ccttctctct	ctgggaacag	2880
tatctggggc	tgccgctgtg	cttctgtctg	gccgtctgca	tctgtctggt	gtgcactttc	2940
ctcgtctgtg	ctctgtctgt	cctcaacccc	tggacggctg	gcctcatagt	gctggctctg	3000
gcgatgatga	cagtggaaat	ctttggtatc	atgggtttcc	tgggcatcaa	gctgagtgcc	3060
atccccgtgg	tgatccttgt	ggcctctgta	ggcattggcg	ttgagttcac	agtccacgtg	3120
gctctgggct	tcctgaccac	ccagggcagc	cggaacctgc	ggccgcccc	tgcccttgag	3180
cacacatttg	cccccgtagc	cgtgggggccc	atctccacat	tgctgggtct	gctcatgctt	3240
gctggttccc	actttgactt	cattgtaagg	tacttctttg	cggcgctgac	agtgtccacg	3300
ctcctggggc	tcctccatgg	actcgtgctg	ctgcctgtgc	tgctgtccat	cctggggccc	3360
ccgcccagag	tgatacagat	gtacaaggaa	agcccagaga	tcctgagtc	accagctcca	3420
cagggaggcg	ggcttaggcc	cgaggagatc	tag			3453

<210> 3
<211> 126
<212> DNA
<213> HUMAN

<400> 3
gagcgcgatgg gcgagtgtct gcagcgcacg ggcaccagtg tcgtactcac atccatcaac 60
aacatggcgg ccttcctcat ggctgccctc gttcccatcc ctgcgctgcg agccttctcc 120
ctacag 126

<210> 4
<211> 119
<212> DNA
<213> HUMAN

<400> 4
gcggccatag tggttggctg cacctttgta gccgtgatgc ttgtcttccc agccatcctc 60
agcctggacc tacggcggcg ccactgccag gcgcttgatg tgctctgctg cttctccag 119

<210> 5
<211> 12886
<212> DNA
<213> HUMAN

<400> 5
cgggtgaatc ccggcgcgcg gccccggacc cgcagctccc tgcactcctc cctcccagcc 60
gctttaaacac ccacacccca cagtctctcc cacgscgcg ccttggcggc cccactgaat 120
ccctacgcgg ggcccagcgg taccgggaga ccgggctagc ctatgggagc gccagataa 180
cgcggttgg gggcgccgcg gcccctatcc ccgcccagc gactcgatcg ccgcccctca 240
gagagctgcc cccgagttac acacccccag ctcgaaaccg agcaccacag gtgagtagag 300
ggggagctgg aagaagggaag agagcggagc caggctctgtc actcgggcct ctgcaagggt 360
tgtgatgtct tgaagtgccg agtgtcatta gatgtctgaa ggcaagttag agccagcacc 420
gcaagcaagt tgtgcgtgtg tgcgggtgtg tctgtgccgg tgtctcctca tcgtctggcc 480
agtgagaatg aatgtctgtg gggtcacctc tgtgtccacc cgacgacagg tgtgtgtaca 540
tatgtatcct gctctcagaa aatgggccta tgccgcccgg cgcggtgact cacgcctgta 600
atcccaaacac tgggaggctg aggcaggcag attacctgag gtcaggagtt cgagaccagc 660
caggccaaca tggggaaact ctgtctctac taaaaataaa aattagcagg gcgtgggtggc 720
ggcgccctgt agtcccaact actcgggagg ctgaggcagg agaatctctt gaacctggga 780
ggcggaggtt gcagtcaagc cgagatcaca ccactgcact ccagccaggg caacagagcg 840
agatgcgtct caaaaaaaaaa aaaaaaaaaa aaaaggagag aaaacaaaaa gaaaagaaaag 900
gaaaataggc ctatgccttc ctccaggtgtg tgcgtgggat ggtgggtgtt acatcttcca 960
agtctgggccc tgtgtctgtg ttgggtgctc ctgtcccaca tccagaaatc aagaagcag 1020
ggctgggcag cagatatata gggtgagaag ggaaggattt catgcattgt tacagtgtatg 1080
cctggtgac ccttctcttt ccattcccaga tcctagctgg gagcctgaag gctccactct 1140
ggcttcgtgc ttacttccag ggctgtctt tctctctggg atgcgggatc cagagacatt 1200
gtggcaaagt gctctttctg ggactgttgg cctttggggc cctggcatta ggtctccgca 1260
tggccattat tgagacaaac ttggaacagc tctgggtaga aggtaagttg tggacactgg 1320
ccatagctgc tcaggatatg tgagcccaag acaagaacgg ggtgaggagc tggctactga 1380
gctctagcag gcctggccct ggggtgggag ggtcacctgc tggccagagt cctacacctg 1440
ggcatctagg aacaccgttc cctcccacaa ccataaatgg acatctgcag atgttagacc 1500
ttgtgaattg ggtgtataga atatttgtct gtgattttgc ttaagcccta gaccacacaa 1560
gaaatccgca gagccacaca accgaggctt ctccacttcc ctggcagagc tccaggaccc 1620
agaagcctgg ggagtgggcc tgggcatga ccaggccctc agggataagg cagggactcc 1680
tctctgcgc acctctccca cttgtctggg gaggcctcag agttagaagc ccccttcgca 1740
agttctgacc cgcttaggcc ctttcccccc cagctttctc tgtttgctct ggcaagtggc 1800
agccgggtga gccaggagct gcattacacc aaggagaagc tgggggagga ggctgcatac 1860
acctctcaga tgctgatata gaccgcacgc caggaggag agaacatcct cacaccgaa 1920
gcacttggcc tccacctcca ggcagccctc actgccagta aagtccaagt atcactctat 1980

gggaaggtga gtctggctga gcccctgagc agctgggggc gaggcgtgct gtggggggttc 2040
tggagtggga atccccttct tctgctgata tccatgccc ctggctattg cagtcctggg 2100
atttgaacaa aatctgctac aagtcaggag tcccccttat tgaaaatgga atgattgagc 2160
gggtaagtgt cctgagaggg agtagaggca gaactttttc tgtagcgtgg gaggactcag 2220
agaccgagca agccccacag cctgcaatct gcccccttaa aactaaggag ggggattgca 2280
gagggcatcc tacaaaggtt gtggggcagg actgacgtgg cccgggggtat ccctggcaga 2340
tgattgagaa gctgtttccg tgcgtgatcc tccccccctt cgactgcttc tgggagggag 2400
ccaaactcca agggggctcc gcctacctgc cgtgagtgc actcctgggg ccctgcttca 2460
tctcccgtcg gggactctcc cagcagaaag gaggggtctg gggaatgagg atgatcaaaa 2520
ccttaccag gtccctaatta cctcccaggc caggaacaga gagcatgggc tcccccaagg 2580
ctctctccac atcctccttc tctttccctc tcaagggaag aagacctgac ttatttacag 2640
aaaactaaac acaaagatct gtaagatctg agcaaaggag aaaaagatcc ccacaaagag 2700
gcttttgctg gggaaattac ctagggtgtt gctaaggcat tgcccaggcc agaaagaaaa 2760
cctgctacag gcatgtgcct gctggttgta tattagaacc aagcacacag cttggtaagg 2820
aactcagtgg ggccctttctg ggccctttct atgtattagg taacctgcc ctgatattcg 2880
tctcagcccc ttgtactctt ctacagctca ctgtagcacc ctggtggggc catgcagcct 2940
ggcagttctg agaagctgag gcttgacac cctccatatt gaaggacaaa tcggcagata 3000
agaggagggg ggggtacagc atggcgcccc agcagcagtt tggagcctgg gttttcgtcc 3060
ctgaccctca ccaactatag gcttttccct cagcgccgc cggtatatcc agtggaacca 3120
cctggatcca cagactgtgc tggaggagct gggtcccttt gcctcccttg agggcttccg 3180
ggagctgcta gacaaggcac aggtgggcca ggccctacgt gggcggccct gtctgcaccc 3240
tgatgacctc cactgcccac ctagtgtccc caaccatcac agcaggcagg tgggttccaa 3300
ccaggtctgc cagggaaagg ctgttttcc tccctttccc tccctcatac tccgtgttc 3360
tgggggagct gactgctctg tgccctgacc cccacttcc tggccattat taccctgtc 3420
ccacagtgc agggcccca tggtccattc ccattcagtt atcctacgga gccctcaagt 3480
ggtatatatg aatccctttt tcccttttca agcctagata aggtggact tctttttttt 3540
ttttttttt agtctcact tgtcaccag gctggagtc agtagttcga tcttggctca 3600
ctgcaacctc ggctcaagca attctcctgc cttagcctcc tgagtagctg ggattacagg 3660
tgcccaccac catgcccggc taatttttat tagcctccca aagtgtctgg attacaggcg 3720
tgagccactg cgctggcca aggtggact ttttatcaaa atagactaat acagggaaac 3780
taagaacaca gcaggtaagc atgaatatca tacctgggtt cccagggttc tttgtggccc 3840
tgcaaatgtg gtactttttt cagaatccgc cagttacacc agctcctccc agaagcctac 3900
ttccaggcct ctgcttcccc ttggggcttc ctgtctgcgg gatactagct gttcactcct 3960
gcagagcact caagaggctc agaatagtta cctacactcc agccctactg agcttcatgg 4020
cagcgtgggt cctggaggtg gaagcccagg gacactcagt tatccacggc cagggccttg 4080
agcattaacc cctcctgttc cctccaggg ctcccaatgt ggctcacgag ctgagtgggg 4140
gctgccatgg cttctccac aaattcatgc actggcagga ggaattgctg ctgggaggca 4200
tggccagaga ccccaagga gagctgctga ggtagggct cctctgggag ttggtgaggg 4260
gactctgttc atgagaacc atactgtaat gccaggcagc tctggcaaaa ggccttcac 4320
atccctcacc aggtgttttg gccagctctg acccctgggt ctcccacacc cccaccaggg 4380
cagaggcct gcagagcacc ttcttgtctg tgagtcccc cgagctgtac gagcatttcc 4440
ggggtgacta tcagacacat gacattggct ggagtggaga gcaggccagc acagtgtctac 4500
aagcttgcca gcggcgcttt gtgcaggctg gtatggacaa ggacaagggg ggtgccctga 4560
ggccattccc tctcctgcc cctcctatc caccctgttt ctccagctgg cccaggaggc 4620
cctgctgag aacgcttccc agcagatcca tgcttctcc tccaccaccc tggatgacat 4680
cctgcatgc ttctctgaag tcagtgtctg ccgtgtgggt ggaggctatc tgctcatggt 4740
gggtcttgca cctggcacct tgccccacc ccacctccaa ccagtgccca cctggggagc 4800
ccctgagact gccctttccc cccacagctg gcctatgcct gtgtgacct gctgcggtgg 4860
gactgcgccc agtcccagg ttccgtggg cttgccggg tactgctggt ggccctggcg 4920
gtggcctcag gccttgggct ctgtgccctg ctgcgcatca ccttcaatgc tgccactacc 4980
caggtacgcc aggactgcag ggcagactca gtgccagtca ccaggcttca cgggtcctca 5040
gctgccgct cctctgcccc tccaggtgct gcccttcttg gctctgggaa tcggcggtga 5100
tgacgtattc ctgctggcg atgccttcac agaggctctg cctggcacc ctctccaggt 5160
ggggccttgt cccccagggc tcatctgagg cagctcagct tactggttaa gagcctcttg 5220
gttcaagtga cccttgggct gctaataaac ctcggtgctt ttgttcttca tctgtaaaac 5280
ggggaataa tagtgctgt tcctaagggt tattgtttgg atcagtgagg taactcaagt 5340
tgaaatgctta gaacagccca tcatacgtac atggtaccca ataaatgcta gccactgtgt 5400
tatgactgcc ccacctctgc accccaagtt cctgagctc cccttctac cactttgaca 5460
cgggccctcc cttgtgacct gagggcaggt cccactctg tctggcagg agcgcatggg 5520
cgagtgtctg cagcgacgg gcaccaggt cgtactcaca tccatcaaca acatggccgc 5580
cttctcatg gctgcctctg ttccatccc tgcgtctgca gccttctccc tacaggcggc 5640

catagtgggt ggctgcacct ttgtagccgt gatgcttgte ttcccagcca tcttcagcct 5700
ggacctacgg cggcgccact gccagcgccct tgatgtgctc tgctgcttct ccagggtactg 5760
cgtgcgcccc agcccccttc tcccgtgacc cagccagcc tgtccctca ccagcatttc 5820
aaggcacaga cctgtcatcc actctctacc tcttccagtc cctgctctgc tcagggtgatt 5880
cagatcctgc cccaggagct gggggacggg acagtaccag tgggcattgc ccacctcact 5940
gccacagttc aagcctttac ccactgtgaa gccagcagcc agcatgtggt caccatcctg 6000
cctccccaag cccacctggt gccccacct tctgaccac tgggctctga gctcttcagc 6060
cctggagggg ccacacggga ccttctaggg caggaggagg agacaaggca gaaggcagcc 6120
tgcaagtccc tgccctgtgc ccgctggaat cttgcccatt tcgcccgcta tcagtttgcc 6180
ccgttgctgc tccagtcaca tgctaaggta agactgggca gagcagggca gagacttagc 6240
atctctgggc ccagaagggc agagagggct tagtccactg cctgaggggc tgggggcagc 6300
cctggggctc ccagcttagt tgctacatcc cgcaggccat cgtgctggtg ctctttggtg 6360
ctcttctggg cctgagcctc tacggagcca ccttggtgca agacggcctg gccctgacgg 6420
atgtggtgcc tcggggcacc aaggagcatg gtgacccagg gtggctttga ctacggccac tcccaacgcg 6480
ccctgtacga ggtggccctg gtgacccagg gtggctttga ctacggccac tcccaacgcg 6540
ccctctttga tctgcaccag cgttccagtt ccctcaaggc ggtgctgccc ccaccggcca 6600
cccaggcacc ccgcacctgg ctgcactatt accgcaactg gctacagggt gagaggcgag 6660
gagacgggca gggaggggtg ctgcaggag aaacggcctg gggccaccag ctaatagaac 6720
cctatcctgg tctccccag gaatccaggc tgcccttgac caggactggg ctctggggc 6780
catcaccgc cactcgaccg caatggctct gaggatgggg cctggccta caagctgctc 6840
atccagactg gagacggcca ggagcttctg gatttcagcc aggttgggag agggctggag 6900
gggtccacta gtacaggggc tgacggccctc ctggggccag gccttcagcc ctctctgct 6960
ctgcagctga ccacaaggaa gctggtggac agagagggac tgattccacc cgagctcttc 7020
tacctggggc tgaccgtgtg ggtgagcagt gacccctgg gtctggcagc ctacagggc 7080
aacttctacc cccacctcc tgaatggctg cagcacaat acgacaccac gggggagaa 7140
tttcgcatg agtcttgggg ggagctcgcc aagagcctca gcctcgcca cacaagccct 7200
gagcctgagg ccctggccac tctgccccgt gctcaccgcc ctgtccctct cctcttctc 7260
ccttccccct ccctccacag tcccggcag ctagcccttg gagtttgccc agttccctt 7320
cctgctgcgt ggccctccaga agactgcaga cttgtggag gccatcgagg gggcccgggc 7380
agcatgcca gaggccggcc aggtgggggt gcacgcctac cccagcggt ccccttctc 7440
cttctgggaa cagtatctgg gcctgcggcg ctgcttctc ctggccgtct gcactctgt 7500
gggtgtgact ttctctgct gtgctctgct gctcctcaac ccctggacgg ctggcctcat 7560
agtgagtgtc tgaggagtg gggacagaga caccaccacc tccctgccc agcctgtcat 7620
ccctcctgcc agggccctc tgtgagccct gtctccctca ggtgctggtc ctggcgatga 7680
tgacagtga actctttggt atcatgggt tcttgggcat caagctgagt gccatccccg 7740
tgggtgatcct tgtggcctt gtaggcatt gcgttgagtt cacagtccac gtggctctgg 7800
tgagcacggg caccggggg agggaccaat cagctgattc agtattcaac acatattgtt 7860
caagccccta ctatgtgcta ggtactattt aagaatttg gctgggtgga cgtggtagct 7920
cattcctgta atccagcac tttgggaggc cgaggcaggt ggatcacctg aggtcaggag 7980
ttcgaacca gcctggccaa catggtgaaa ccctgtctt actaaaaata caaaaaatta 8040
gccaggcgtg gtggcacatg ccagtaatcc cagctacttt ggaggctgag gcagaattgc 8100
ttgaacctgg gaggcgaagg ttgcagttag ctgagatcgt gccattgcac tccagctgg 8160
gcaacaagag tgaactctc cgtctcaaaa aaaaaaattt gggctgggca 8220
cagtggctca tgcctgtaat tgggatgatg ctggggcatt ttgggaggcc aaggcaggcg 8280
gatccccga agtcaggagt tcaagaccag cctggccaac attgcaaaac cccgtctcta 8340
ctgaaaatac aaaaattagc tgggcgtggt ggctcatgcc tgtaatccca gctactcagg 8400
aggctgaggc aggagaatta cttgaaccca ggaggcggag gttacagtga gctgagatca 8460
catcactgta ctccagcctg ggcaaaagag caagattcaa tcttcaaaaa aagaatttg 8520
aaaataaaaa taaaaagaat acgggatata atagcaatac aggttgggtt acgcctgtaa 8640
cttatattct agggatagag atagacaata agggctgggt gatgtggctt tcaagaccag 8700
tcccagcact ttgggaagcc gaggtgggca catcacttga ggtcaggagt tgggtgtggg 8760
cctggccgac atggtgaaac cccatctcta ctaaaaatat gaaaattagc cttgaacccg 8820
ggtgcatgcc tgtaatccca gctacttggg aggtcagggc aggagaatca actccagcct gggagaccga 8880
ggagggtgga ggttgccatg agccgagacc atgccactgc actccagcct ccccaaagt 8940
gcaaaactcc atctcaaaaa aagaaaaaaa aacctcagcc tcccaaagt acagggctc 9000
aggcatgagc caccgtgccc gggcttgttt gttttaagag actctgtc acatccatac 9060
ccaggctgta ccctatttct tacgtgtctc tgtgtctaag ctaccaaac ccatccatac 9120
agtgttacct aaaaggactc acgcatgctc cctgtacagt tcccatgaac ttagttcacc 9180
cagcctttgg tgggccaagg tatcgttgag ctctgaagac agatacagct cgggacatgg 9240
ctgagctggc catgactggc agaggagcag ctccaggacc actctgtttt cctagggctt 9300
cctgaccacc cagggcagcc ggaacctgcg ggccgcccac gcccttgagc acacatttgc

ccccgtgacc	gatggggcca	tctccacatt	gctgggtctg	ctcatgcttg	ctgggtccca	9360
ctttgacttc	attgtaaggt	agggagggct	cggggcaggg	aggcagggct	caggacaggc	9420
ctgggctgac	tccccccaca	ccctacccct	aggtacttct	ttgcggcgct	gacagtgttc	9480
acgtctctgg	gctctctcca	tggactcgtg	ctgctgcctg	tgctgtgtgc	catcctgggc	9540
ccgccgccag	aggtgaccac	accctcggca	ccatccctct	actcccagcc	caagggacgg	9600
ggtagggaga	ggcaagggaa	gggacagagc	cctgtggccc	acagacaggt	acctcccaa	9660
caggtgccac	cagctgaagg	tggcagcctc	ctcctttccc	cagacaccat	gttctgtccc	9720
ctcagccctc	ctggcttctt	catgggaccc	accttagact	tttaggatcc	agaacaaggt	9780
gcagggtttg	ccccaggcct	caacatcctg	tgcctgcca	gctctcatat	cctgtctggag	9840
accaacaagg	gccccagctt	cccaacagtc	atggtaatcc	ccagcgagat	gctaaagggg	9900
acgggagccc	cagggggccc	tgggcttact	ggggctgggt	tctccccaca	ggtgatacag	9960
atgtacaagg	aaagcccaga	gaccttgagt	gcacagctc	cacagggagg	cgggcttagg	10020
tggggggcat	cctcctccct	gccccagagc	tttgccagag	tgactacctc	catgaccgtg	10080
gccatgccacc	cacccccccct	gcctggtgcc	tacatccatc	cagccccctga	tgagccccct	10140
tgggtccctg	ctgtcactag	ctctggcaac	ctcagttcca	ggggaccagg	tccagccact	10200
gggtgaaaga	gcagctgaag	cacagagacc	atgtgtgggg	cgtgtggggg	cactgggaag	10260
cactgggtct	ggtgttagac	gcaggatgga	cccctggagg	gctctgctgc	tgctgcatcc	10320
cctctcccga	cccagctgtc	atgggcctcc	ctgatatacca	tacagaacag	ccaccgattt	10380
gcacatccag	gctgtgtga	gcctgtatct	gtgtcacttg	agagtgaag	ctggcacttg	10440
gggctgcagt	gcagcctctg	ccccctccc	accaccacc	actgcctgcc	cagctgacca	10500
agcctgaggg	accctccagc	acccttccgt	ctggtgactc	ctgggcaggg	tctccatata	10560
cctgccacc	tcctaccaca	tccattatth	atatgaaaat	gtctatthtt	gtagtataca	10620
tacatgttag	ctatgatgaa	agthttatth	thtaagaat	gaaatatatt	ctatgtgaag	10680
ctatgatgaa	agthttatth	thtaagaat	gaaatatatt	ctatgtgaac	taatctcgaa	10740
agthttatth	thtaagaat	gaaatatatt	ctatgtgtgc	aagtgaacat	tagcttcagt	10800
tgctthttth	tggacagagt	ggggagthtt	caagtgaaca	ttagctattg	gaaggagctt	10860
ctctggtgcc	aggacctgag	gtattagctt	ctctagthtt	gggtggaaaa	gacccagat	10920
tctggtatth	gtcatatac	ttggtaacat	catctggatt	aagtgttac	tatacaaaac	10980
gataacaaat	thttgtgtgt	tgaatccta	ctgggttcaa	tctggagacc	gagagcagaa	11040
aaaaaagaac	cccactgtgt	ggctthcaga	gccaccatat	tccagcctgc	ccgtctctcc	11100
agactcacct	ccacctacct	gcttcacccg	cacgggaac	ggcaaggcag	aggggcaaa	11160
ccatgcagca	ggtggaaggc	gaggtggagg	cagatcagga	aagcagccag	ttgaagcaga	11220
gagaggtcaa	cagggctctg	ggagcttctc	aggaggtttg	tggacccagg	gaaaggagcc	11280
aggttccaga	gcaacctcca	aggcaaaggc	ctctgtaagt	tggttgtcct	gacagccag	11340
aggtgtctth	ggccagtcag	ccagtggatc	agttgcggga	actgctcaga	aactgaggtg	11400
ctagcagtht	gtgaggacac	agcgtaaagt	gtttgttctg	tgaagttga	acagctccac	11460
taagcagagg	ccttgaagag	tggccacagc	cctggaatag	agcacagagc	ctcacctaga	11520
ggcgtgggga	ggtttgcaac	tgcccttccc	cagccatagc	ttaggaccca	tagtctagtt	11580
cacatagacc	ctgggtccca	accacccact	caccaggaat	gatccccccc	caggaacaat	11640
gcgttctcac	atcccccccc	acctggacaa	aggccaggaa	atcatgttct	gacccaaaaga	11700
tacaacaaca	aaaacaacaa	caacaaaaaa	cgctatttgc	aattgaatcc	acgctaaaa	11760
gcctaaaaag	ctcaagagaa	gcgggtagtt	ggcagagaac	ctagagtagg	gggtgcaacc	11820
agcaggccca	agggaggag	gctgcatttg	ggtccagcag	tgthttgggtc	accaagaagg	11880
gccttctagg	tggagcagag	agagctcacc	aggccagaat	agtgcacagg	gggtcagccc	11940
tcagtgccac	ttaccagcgg	agtaaccttg	ggcaagttag	ccagcctcac	taagcctccc	12000
catcttcatc	thttccaggcc	cgaggagatc	tagcctctgc	ctcccccccc	agcaccctct	12060
catcagacac	aaggagcgcc	actgtctgga	caggctgaat	tggthttcgg	gtccctaat	12120
tctcatacgc	cattccctct	gcctagaaca	ctthttcacc	tccccctgat	gtgaccccat	12180
atcaccttcc	gaggtgaatt	ggatcggtg	ccatctcctc	caggaggggt	gggtcgtgc	12240
ctcctgtgag	gtcccagthc	ccctgagthc	ctgtgcccgt	ctgtthtccc	gtccctctct	12300
ctaagcccg	agccttactg	cggttaagga	cgccgggaga	ggaccttaac	ccctgggacg	12360
aacaccagct	ccgcaaagga	ctccgcaccc	ggcgccgccc	acgggggtgcg	ggtcccagga	12420
ggaccagcag	agaggagcat	aggagagcaa	aggagatcag	tgacctatgg	ctthccccgg	12480
ggcgcggaac	agcccgagc	cgctgtgat	ttgcataccc	atggtgcacc	acgaaaagat	12540
acctcaaga	tgcttgcact	ccctctgtgc	gcgcatttct	gcactgttht	agagcatgat	12600
gcctcttaca	cgcatctgtg	tgcataaaact	acatataggg	agthcgtacc	acgcaggcat	12660
ccaacaacca	taagtgtgtt	aagtgttagt	tctccctgcg	agthtcgaag	cggaagthc	12720
gaatatactc	gggtthtctct	tcaaagcgca	thaatcttht	gcctthtact	aaagatttht	12780
gtggagagaa	agthttagt	thttattcaa	thtttgagg	ctctthttht	cctgaggcta	12840
cattthttaag	tattaaaagt	taggcaacta	caaaaaaaa	aaaaaa		12886

1

INTERNATIONAL SEARCH REPORT

International application No.

PCT/SE 99/01784

A. CLASSIFICATION OF SUBJECT MATTER

IPC7: A61K 48/00, C07K 14/705, A61K 38/17, G01N 33/68
According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)

IPC7: A61K, C07K, G01N

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

SE,DK,FI,NO classes as above

Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
P,X	WO 9929854 A1 (ONTOGENY, INC.), 17 June 1999 (17.06.99), page 73 - page 77, claim 3 --	1-12,16-22
P,X	EP 0879888 A2 (SMITHKLINE BEECHAM PLC), 25 November 1998 (25.11.98), claim 9b --	1-12,16-22
P,X	Journal of Clinical Investigation Online, Volume 95, No 23, November 1998, David Carpenter et al., "Characterization of two patched receptors for the vertebrate hedgehog protein family", pages 13630-13634 --	1-12,16-22

☒ Further documents are listed in the continuation of Box C.

☒ See patent family annex.

* Special categories of cited documents:

"A" document defining the general state of the art which is not considered to be of particular relevance

"E" earlier document but published on or after the international filing date

"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)

"O" document referring to an oral disclosure, use, exhibition or other means

"P" document published prior to the international filing date but later than the priority date claimed

"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention

"X" document of particular relevance: the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone

"Y" document of particular relevance: the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art

"&" document member of the same patent family

Date of the actual completion of the international search

1 February 2000

Date of mailing of the international search report

14 -02- 2000

Name and mailing address of the ISA/
Swedish Patent Office
Box 5055, S-102 42 STOCKHOLM
Facsimile No. +46 8 666 02 86

Authorized officer

Patrick Andersson/AE
Telephone No. +46 8 782 25 00

INTERNATIONAL SEARCH REPORT

International application No.

PCT/SE 99/01784

C (Continuation). DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	Nature Genetics, Volume 18, February 1998, Motoyama J. et al., "Ptch2, a second mouse Patched gene is co-expressed with Sonic hedgehog", page 104 - page 106, figure 1, EMBL Acc no AB000847 --	1-12,16-22
A	FEBS Letters, Volume 410, 1997, Takashi Takabatake et al., "Hedgehog and patched gene expression in adult ocular tissues" page 485 - page 489 --	1-12,16-22
A	WO 9745541 A2 (THE LELAND S. STANFORD JUNIOR UNIVERSITY), 4 December 1997 (04.12.97) -- -----	1-12,16-22

INTERNATIONAL SEARCH REPORT

International application No.

PCT/SE99/01784

Box I Observations where certain claims were found unsearchable (Continuation of Item 1 of first sheet)

This international search report has not been established in respect of certain claims under Article 17(2)(a) for the following reasons:

1. ☒ Claims Nos.: claims 12 partially and claim 21 completely because they relate to subject matter not required to be searched by this Authority, namely:

See next sheet.

2. ☒ Claims Nos.: 13-15 because they relate to parts of the international application that do not comply with the prescribed requirements to such an extent that no meaningful international search can be carried out, specifically:

See next sheet.

3. ☐ Claims Nos.: because they are dependent claims and are not drafted in accordance with the second and third sentences of Rule 6.4(a).

Box II Observations where unity of invention is lacking (Continuation of Item 2 of first sheet)

This International Searching Authority found multiple inventions in this international application, as follows:

1. ☐ As all required additional search fees were timely paid by the applicant, this international search report covers all searchable claims.
2. ☐ As all searchable claims could be searched without effort justifying an additional fee, this Authority did not invite payment of any additional fee.
3. ☐ As only some of the required additional search fees were timely paid by the applicant, this international search report covers only those claims for which fees were paid, specifically claims Nos.:
4. ☐ No required additional search fees were timely paid by the applicant. Consequently, this international search report is restricted to the invention first mentioned in the claims; it is covered by claims Nos.:

Remark on Protest

☐

The additional search fees were accompanied by the applicant's protest.

☐

No protest accompanied the payment of additional search fees.

INTERNATIONAL SEARCH REPORT

International application No.

PCT/SE99/01784

Box I. 1.

Claims 12 partially 21 completely relate to methods of treatment of the human or animal body by therapy/ diagnostic methods practised on the human or animal body See PCT Rule. 39.1.(iv). Nevertheless, a search has been executed for these claims. The search has been based on the alleged effects of the compounds.

Box I. 2.

Claim 13 relates to a screening method, it is not clear from this claim or the description how the polypeptide is used to achieve the screening. Claim 14 relates to a method of synthesis a modified drug, it is not clear from the description or the claim how the protein is used. Claim 15 relates to a drug identified by any of the methods in 13 or 14. Thus, claims 13-15 is considered to fail to comply with PCT-article 6 and have not been searched. Moreover, claim 15 concerns a drug per se, the drug may be a known substance. The description or the claims does not give any indication on that any unknown substance(s) that are preferred. For this reason, a search or a statement concerning novelty and an inventive step can not be made for claim 15.

Claims 12 partially 21 completely relate to methods of treatment of the human or animal body by therapy/ diagnostic methods practised on the human or animal body See PCT Rule. 39.1.(iv). Nevertheless, a search has been executed for these claims. The search has been based on the alleged effects of the compounds.

Claim 13 relates to a screening method, it is not clear from this claim or the description how the polypeptide is used to achieve the screening. Claim 14 relates to a method of synthesis a modified drug, it is not clear from the description or the claim how the protein is used. Claim 15 relates to a drug identified by any of the methods in 13 or 14. Thus, claims 13-15 is considered to fail to comply with PCT-article 6 and have not been searched. Moreover, claim 15 concerns a drug per se, the drug may be a known substance. The description or the claims does not give any indication on that any unknown substance(s) that are preferred. For this reason, a search or a statement concerning novelty and an inventive step can not be made for claim 15.

INTERNATIONAL SEARCH REPORT
Information on patent family members

02/12/99

International application No.
PCT/SE 99/01784

Patent document cited in search report			Publication date	Patent family member(s)		Publication date
WO	9929854	A1	17/06/99	AU	1716599 A	28/06/99

EP	0879888	A2	25/11/98	CA	2232808 A	23/11/98
				GB	9710752 D	00/00/00
				JP	11075874 A	23/03/99
				GB	9805954 D	00/00/00

WO	9745541	A2	04/12/97	AU	3227497 A	05/01/98
